Geomorphology : Theoretical and Applied

GEOMORPHOLOGY
Theoretical and Applied

Dr. Udhav Eknath Chavan

Chandralok Prakashan
KANPUR-208 021 (INDIA)

Geomorphology : Theoretical and Applied

First Published : 2022
ISBN : 978-93-93561-15-2

Author : Dr. Udhav Eknath Chavan

Published by
CHANDRALOK PRAKASHAN
132, 'Shivram Kripa', Mayur Park,
Basant Vihar, Kanpur - 208 021
Ph : 0512-2634444, 0512-2634242, Fax : 0512-2634444
Mob. : 09415125867, 9506294444, 09415200584
E-mail : chandralok.prakashan@gmail.com
info@chandralokprakashan.com
visit us at : www.chandralokprakashan.com

PRINTED IN INDIA

Printed at Deepak Offset Press, Delhi.

Preface

Geomorphology is the study of the nature and history of landforms and the processes which create them. Initially, the subject was committed to unravelling the history of landform development, but to this evolutionary approach has been added a drive to understand the way in which geomorphological processes operate.

Geomorphology is the scientific study of landforms and the processes that shape them. Geomorphologists seek to understand why landscapes look the way they do, to understand landform history and dynamics, and to predict future changes through a combination of field observations, physical experiments, and numerical modeling. Geomorphology is practiced within geography, geology, geodesy, engineering geology, archaeology, and geotechnical engineering, and this broad base of interest contributes to a wide variety of research styles and interests within the field.

Geomorphology is the study of Earth's landforms created by mostly physical processes, including physical or chemical changes and those changes influenced by biological processes, including land use. Physical geographers apply geomorphological principals to study how landforms have changed in the past, but increasingly such principals are important for modern applications. Over long geological timespans, plate tectonics have shaped continents. Earthquakes and volcanic activity represent active change that relate to plate tectonic movements. Fluvial, or those involving water, change is among the most significant physical factors that shape the Earth at generally small scales.

Geomorphology is the study of the features that make up the earth's surface and their relationship to the underlying geology.

A geomorphological study will provides a conceptual picture of coastal processes and the potential behaviour of the coastal system. This includes taking into account changes in the bedrock composition that could affect the potential rate of future coastal evolution. The results tend to be qualitative, rather than quantitative. This chapter starts with a description of how a sediment budget may be used to provide a view about future beach levels in front of a coastal structure. The section then moves onto describe useful projects that have has a significant geomorphological component, namely Futurecoast and Eurosion and introduces the concept of the coastal tract as a way of approaching very long term coastal evolution.

Geomorphology has traditionally focused on the study of landforms and on the processes involved in their formation. Applied geomorphology is the practical application of this study to a range of environmental issues, both in terms of current problems and of future prediction. Applied geomorphology provides a strategic tool for informed decision-making in public policy development and in environmental resource management. Key areas of application include specific environmental settings, such as the coastal zone or dryland environments; the impacts of land use and management practice on Earth surface processes; and areas susceptible to natural hazards.

It is hoped that the scholars working in the field of applied geomorphology, geology, settlement and land use planning, environmental management, civil engineering and other allied disciplines will benefit from this work and will be quick to follow the methods and techniques used therein to achieve their objectives.

—Dr. Udhav Eknath Chavan

Contents

1

Historical and Process Geomorphology

Geomorphology is the scientific study of landforms and the processes that shape them. Geomorphologists seek to understand why landscapes look the way they do, to understand landform history and dynamics, and to predict future changes through a combination of field observations, physical experiments, and numerical modeling. Geomorphology is practiced within geography, geology, geodesy, engineering geology, archaeology, and geotechnical engineering, and this broad base of interest contributes to a wide variety of research styles and interests within the field.

The Historical Approach

Historical studies attempts to deduce from the erosion and depositional features of the landscape evidence relating to the sequence of historical events. For example, tectonic, sea level, climatic through which it has passed. Such studies explain the existing landform assemblages as a combination of effects resulting from the vicissitudes (changes) through which it has passed. Historical explanation is reserved for landforms whose features have evolved slowly and which bear witness to the super imposed effects of climatic and tectonic changes. It relies retro diction, the derivation of chronology of a sequence of past landscape forming events. Historical approach comprise of two related approaches namely the cyclic approach and the denudation chronology approach.

1. Cyclic Approach : the cyclic approach was initiated and established by W.M Davis (1850- !9340), who summarized the basic thesis of his work in the phrase Landforms are a function of structure process and stage. Although he concentrated on the sequence of events in explaining the evolution of landforms, this evolutionary sequence he termed the cycle based on his conceptualization that landforms like the human life pass through the stages of youth, maturity and old age. A major feature of any cycle is that the end position should be similar to the initial one. Thus, the end product of Davisian cycle of erosion, the peneplain is about the same as the initial land surface (peneplain means a flat or nearly flat landscape).
2. The Denudation Chronology Approach - the mainstream of geomorphological thought particularly in Britain from about 1930- 1950 was denudation chronology. This was the deliberated attempt to make use of erosion surfaces at different attitudes to reconstruct the geomorphological history of a region. An erosion surface is an extensive flat area known to have been produced by erosion and representing the product of a cycle denudation chronology is thus, not concerned about the origin of just one landform, but an assemblage of landforms or surface. The main roles of denudation chronology (has its main roles) are the identification, dating and interpretation of Planation surfaces. In addition, it has the subsidiary aim of studying the way in which the drainage system of an area has evolved. However, denudation chronology involves the absolute or relative dating of erosional and depositional events occurring under the influence of tectonic, eustatic, climatic, or other variations. The techniques used to identify erosion surfaces comprising a combination of field observation and analysis of contoured map.

Ancient geomorphology

The study of landforms and the evolution of the Earth's surface can be dated back to scholars of Classical Greece. Herodotus argued from observations of soils that the Nile delta was actively growing into the Mediterranean Sea, and estimated its age. Aristotle speculated that due to sediment transport into the sea, eventually those seas would fill while the land lowered. He claimed that this would mean that land and water would eventually swap places, whereupon the process would begin again in an endless cycle.

Another early theory of geomorphology was devised by the polymath Chinese scientist and statesman Shen Kuo (1031–1095). This was based on his observation of marine fossil shells in a geological stratum of a mountain hundreds of miles from the Pacific Ocean. Noticing bivalve shells running in a horizontal span along the cut section of a cliffside, he theorized that the cliff was once the pre-historic location of a seashore that had shifted hundreds of miles over the centuries.

He inferred that the land was reshaped and formed by soil erosion of the mountains and by deposition of silt, after observing strange natural erosions of the Taihang Mountains and the Yandang Mountain near Wenzhou. Furthermore, he promoted the theory of gradual climate change over centuries of time once ancient petrified bamboos were found to be preserved underground in the dry, northern climate zone of *Yanzhou*, which is now modern day Yan'an, Shaanxi province.

Early modern geomorphology

The term geomorphology seems to have been first used by Laumann in an 1858 work written in German. Keith Tinkler has suggested that the word came into general use in English, German and French after John Wesley Powell and W. J. McGee used it during the International Geological Conference of 1891. John Edward Marr in his The Scientific Study of Scenery considered his book as, 'an Introductory Treatise on Geomorphology, a subject which has sprung from the union of Geology and Geography'.

An early popular geomorphic model was the *geographical cycle* or *cycle of erosion* model of broad-scale landscape evolution developed by William Morris Davis between 1884 and 1899. It was an elaboration of the uniformitarianism theory that had first been proposed by James Hutton (1726–1797). With regard to valley forms, for example, uniformitarianism posited a sequence in which a river runs through a flat terrain, gradually carving an increasingly deep valley, until the side valleys eventually erode, flattening the terrain again, though at a lower elevation. It was thought that tectonic uplift could then start the cycle over. In the decades following Davis's development of this idea, many of those studying geomorphology sought to fit their findings into this framework, known today as "Davisian". Davis's ideas are of historical importance, but have been largely superseded today, mainly due to their lack of predictive power and qualitative nature.

In the 1920s, Walther Penck developed an alternative model to Davis's. Penck thought that landform evolution was better described as an alternation between ongoing processes of uplift and denudation, as opposed to Davis's model of a single uplift followed by decay. He also emphasised that in many landscapes slope evolution occurs by backwearing of rocks, not by Davisian-style surface lowering, and his science tended to emphasise surface process over understanding in detail the surface history of a given locality. Penck was German, and during his lifetime his ideas were at times rejected vigorously by the English-speaking geomorphology community. His early death, Davis' dislike for his work, and his at-times-confusing writing style likely all contributed to this rejection.

Both Davis and Penck were trying to place the study of the evolution of the Earth's surface on a more generalized, globally relevant footing than it had been previously. In the early 19th century, authors – especially in Europe – had tended to attribute the form of landscapes to local climate, and in particular to the specific effects of glaciation and periglacial processes. In contrast, both Davis and Penck were seeking to emphasize the importance of evolution of landscapes through time and the generality of the

Earth's surface processes across different landscapes under different conditions.

During the early 1900s, the study of regional-scale geomorphology was termed "physiography". Physiography later was considered to be a contraction of "*physi*cal" and "ge*ography*", and therefore synonymous with physical geography, and the concept became embroiled in controversy surrounding the appropriate concerns of that discipline.

Some geomorphologists held to a geological basis for physiography and emphasized a concept of physiographic regions while a conflicting trend among geographers was to equate physiography with "pure morphology", separated from its geological heritage.

In the period following World War II, the emergence of process, climatic, and quantitative studies led to a preference by many earth scientists for the term "geomorphology" in order to suggest an analytical approach to landscapes rather than a descriptive one.

Process Geomorphology

Wind, flowing water, and moving ice — all bring about changes in Earth's surface, as do the plants, animals, and humans. Powered by the Sun, each of these 'agents' of geomorphic change has its unique set of processes through which it transforms landscapes and leaves its signature. The process–landform relationship is the basic foundation for the quantitative assessment of geomorphic systems. Formulation of mathematical models through rational deduction and empirical analysis of observed data, to relate energy, mass, and time is the ultimate goal of process geomorphology. The outcome of such analyses forms the key component in managing geomorphic hazards.

Quantitative and process geomorphology

Geomorphology was started to be put on a solid quantitative footing in the middle of the 20th century. Following the early work of Grove Karl Gilbert around the turn of the 20th century, a group of mainly American natural scientists, geologists and hydraulic engineers including William Walden Rubey, Ralph Alger Bagnold, Hans Albert Einstein, Frank Ahnert, John Hack, Luna Leopold, A. Shields, Thomas Maddock, Arthur Strahler, Stanley Schumm, and Ronald Shreve began to research the form of landscape elements such as rivers and hillslopes by taking systematic, direct, quantitative measurements of aspects of them and investigating the scaling of these measurements..

These methods began to allow prediction of the past and future behavior of landscapes from present observations, and were later to develop into the modern trend of a highly quantitative approach to geomorphic problems.

Many groundbreaking and widely cited early geomorphology studies appeared in the Bulletin of the Geological Society of America, and received only few citations prior to 2000 (they are examples of "sleeping beauties") when a marked increase in quantitative geomorphology research occurred.

Quantitative geomorphology can involve fluid dynamics and solid mechanics, geomorphometry, laboratory studies, field measurements, theoretical work, and full landscape evolution modeling. These approaches are used to understand weathering and the formation of soils, sediment transport, landscape change, and the interactions between climate, tectonics, erosion, and deposition.

In Sweden Filip Hjulström's doctoral thesis, "The River Fyris" (1935), contained one of the first quantitative studies of geomorphological processes ever published. His students followed in the same vein, making quantitative studies of mass transport (Anders Rapp), fluvial transport (Åke Sundborg), delta deposition (Valter Axelsson), and coastal processes (John O. Norrman). This developed into "the Uppsala School of Physical Geography".

HUTTON TO HORTON

The first use of the word geomorphology was likely to be in the German language when it appeared in Laumann's 1858 work.

Keith Tinkler has suggested that the word came into general use in English, German and French after John Wesley Powell and W. J. McGee used it in the International Geological Conference of 1891.

An early popular geomorphic model was the *geographical cycle* or the *cycle of erosion*, developed by William Morris Davis between 1884 and 1899. The cycle was inspired by theories of uniformitarianism first formulated by James Hutton (1726–1797). Concerning valley forms, uniformitarianism depicted the cycle as a sequence in which a river cuts a valley more and more deeply, but then erosion of side valleys eventually flatten the terrain again, to a lower elevation. Tectonic uplift could start the cycle over. Many studies in geomorphology in the decades following Davis' development of his theories sought to fit their ideas into this framework for broad scale landscape evolution, and are often today termed "Davisian". Davis' ideas have largely been superseded today, mainly due to their lack of predictive power and qualitative nature, but he remains an extremely important figure in the history of the subject.

In the 1920s, Walther Penck developed an alternative model to Davis', believing that landform evolution was better described as a balance between ongoing processes of uplift and denudation, rather than Davis' single uplift followed by decay. However, due to his relatively young death, disputes with Davis and a lack of English translation of his work his ideas were not widely recognised for many years.

These authors were both attempting to place the study of the evolution of the Earth's surface on a more generalized, globally relevant footing than had existed before. In the earlier parts of the 19th century, authors-especially in Europe-had tended to attribute the form of landscape to local climate, and in particular to the specific effects of glaciation and periglacial processes. In contrast, both Davis and Penck were seeking to emphasize the importance of evolution of landscapes through time and the generality of Earth surface processes across different landscapes under different conditions.

The papers by Strahler (1952) and Chorley (1962) strongly advocated the adoption of a "dynamic" as opposed to an "historical" approach to geomorphology. The opinion of some later workers—notably Simpson (1963) and Mayr (1982) —is, however, that any advance in the historical natural sciences depends upon the combined appreciation of immanent and configurational elements (Simpson's terminology); and the view that events may have an essential historical or timebound component is now accepted even in "experimental" sciences such as chemistry (Prigogine, 1978). In the light of these contrasting approaches to earth science, an attempt is made to analyse the mjor lines of thought concerning change, progression and equilibrium in the work of six leading precursors of modern geomorphology: James Hutton, Charles Lyell, Charles Darwin, James Dwight Dana, Grove Karl Gilbert and Robert E. Horton. Despite their perceived general adherence to the Uniformitarian tradition, it is suggested that the work of the six reveals two contrasting attitudes to ideas of change and of equilibrium, It is argued that those authors — Lyell, Dana, Horton—who are primarily concerned to demonstrate that the present state of the earth is in some sense the normal or optimum, tend at the same time to accept the existence or desirability of some equilibrium state and, paradoxically, to overstate the role of "unusual", "cataclysmic" or "catastrophic" events in creating and sustaining this equilibrium. The views of Horton, in particular, lend themselves to the description "punctuated equilibrium". In contrast, it is contended that Hutton, Darwin and Gilbert have no ideological commitment to the present state of the earth as anything other than one moment in time. Their ideas are considered to focus upon the entire sequence of changes which may be inferred to create the phenomena viewed at any time or place.

Quantification

Quantification began to sweep geomorphology after the publication of R. E. Horton´s visionary studies of drainage basin analysis. Some attempts at quantification were decidedly innovative. In the case of M. A. Melton, the work was so ahead of its time that only a few geomorphologists appreciated its implications. The extensive work on drainage basin and hillslope quantifications by Strahler and his students inspired a flowering of geomorphic research in the 1960s.

Similarly, the detailed studies of small-scale fluvial processes by Luna Leopold and colleagues at the U.S. Geological Survey also led to an abundance of related studies. Such process studies were decidedly advanced by technological developments that allowed for relatively easy measurement and long-term monitoring of processes in the field. Morphometric studies proved amenable to automatic data processing procedures by computer. Quantification has been described as a revolution in geomorphology. Although its use certainly superseded the qualitative approach of William Morris Davis, it is clear that quantification never constituted a revolution in the accepted sense of scientific philosophy. Quantification is a tool of study, one that indeed adds great power to the simplification of complexity. Nevertheless, it remains a mere technique, not a fundamental framework of thought.

One exciting aspect of quantification in mega-geomorphology derives from the ability of computing systems to handle the immense data sets necessary to describe terrain. The manipulation of these very large data sets will generate new and interesting research questions.

HORTON TO STRAHLER-HACK

While Penck and Davis and their followers were writing and studying primarily in Western Europe, another, largely separate, school of geomorphology was developed in the United States in the middle years of the 20th century.

Following the early trailblazing work of Grove Karl Gilbert around the turn of the 20th century, a group of natural scientists, geologists and hydraulic engineers including Ralph Alger Bagnold, John Hack, Luna Leopold, Thomas Maddock and Arthur Strahler began to research the form of landscape elements such as rivers and hillslopes by taking systematic, direct, quantitative measurements of aspects of them and investigating the scaling of these measurements.

These methods began to allow prediction of the past and future behavior of landscapes from present observations, and were later to develop into what the modern trend of a highly quantitative approach to geomorphic problems. Quantitative geomorphology can involve fluid dynamics and solid mechanics, geomorphometry, laboratory studies, field measurements, theoretical work, and full landscape evolution modeling. These approaches are used to understand weathering and the formation of soils, sediment transport, landscape change, and the interactions between climate, tectonics, erosion, and deposition.

The Horton-Strahler scheme for stream network classification was introduced into hydrological practice in the middle of the last century. It found a significant application in the process of defining the Geomorphologic Instantaneous Unit Hydrograph (GIUH) as the method aiming to relate geomorphological characteristics of a basin to basin response in the form of unit hydrograph, or direct runoff hydrograph. The paper analyzes the problems related to its implementation in the GIUH method, i.e. to determination of initial and transition probabilities, which, in turn, are used to define the probability distribution function for water flow paths in the basin

THE PARADIGM CHANGE

The acceptance of new ideas into the mainstream of geomorphological education is illustrated from the development of theories dealing with Earth history, glaciation, uniform flow, mass movement, continental mobility, cyclic erosion, and drainage networks. The lag between the conception of new ideas and their incorporation into mainstream texts has varied from negligible to more than 200 years. On one hand, despite its then untestable assumptions, the Davisian cycle of erosion gained rapid favor as the dominant paradigm of the early 20th century before it was found wanting. In contrast, concepts of uniform flow and slope stability, confirmed in the 18th century, waited almost 200 years for incorporation into geomorphology texts sensu stricto, although they had long been available in books on hydraulics and soil mechanics. Continental mobilism had a wild ride, culminating in the eventual acceptance of the plate-tectonics paradigm in the later 20th century.

At least since the time of Gilbert and Davis, geomorphologists have pursued and employed theory-laden observational techniques. From the Davisian geographical cycle which is basically descriptive, the quantitative-dynamic approach to landform studies and hydraulic geometry, to the systematic geomorphology today, there are full of paradigms, principles and basic concepts. Full recognition and understanding of these paradigms are essential for developing a unified approach to the science of geomorphology.

2

Theories, Techniques and Fieldwork (Including Field Experiments) in Geomorphology

THEORIES OF LANDFORM EVOLUTION

U.S. geomorphology experienced spectacular growth from about 1890 to about 1950. American academics recognized W.M. Davis (1850–1934) as the leading specialist in the field during this period. Davis developed the technique of "explanatory description" of landscapes. He discouraged experimentation in processes of landscape evolution, because he felt the processes were too complex and landform types too varied. His methodology focused instead on describing three controlling factors of landscape change—structure, process, and time.

Davis devised one of the earliest models of landscape evolution—the geographical cycle. The model posited that landscapes evolve through a sequential series of stages; each stage possesses an indicative set of landforms. Purely descriptive in origin, the observer was supposed to use the model to determine the stage of development of a particular landscape. Followers of Davis conducted few experimental studies on processes, for his descriptive system was deceptively adequate. Intuitive reasoning outpaced investigational proof during these early years. World War II introduced improvements in aerial photography and new kinds of geomorphologic analysis. Accurate quantitative interpretation of beaches and coasts using aerial photographs were vital to military strategy during the war. Beach studies carried over after the war and led the trend in quantification and experimental verification.

A concern for streams, the main suppliers of sand to beaches, was a logical extension of shoreline studies. Thus, the dynamics of weathering rates, stream erosion, and slope development quickly became a major focus of the new approach. By the late 1960s, quantitative models had all but supplanted Davisian-type descriptive models. In subsequent decades, the role of climate in the genesis of landforms received more interest. Additionally, new technologies involving the study of the ocean floor validated the unifying theory of plate tectonics. Theoretical geomorphology in the United States has used the systems theory approach and has focused on developing quantitative models that replicate and predict geological processes. Geographic information systems, remote sensing, global positioning systems, and isotope dating have accelerated this trend.

GEOGRAPHERS AND THE THEORY

Fieldwork, especially in "foreign" lands, is one of the time-honoured hallmarks of professional geography, yet as Felix Driver and others have noted, "little attention has been paid to the specifically geographical dimensions of field-work [sic] or to its history within the discipline." In this chapter I will explore pieces of this history through the practice of return visitation to particular places, with an emphasis on the record of North American geographers' engagement with Latin America. In turn, Carl Sauer's Mexican research, together with that of some of his students and others inspired by his example, appears to be the most tightly woven net of repeat visitation. Here, one finds multiple strands of crisscrossing lines of reconnaissance and research, and complicated proxy relations involving individuals from later generations pursuing topics initiated by individuals from earlier ones.

Space here does not permit more than sketching the outlines of this history, but it does illustrate some of the salient aspects of repeat visitation, and points to a neglected topic in the history of geography. Viewed from the broad perspective of geography's historical record, the practice of return or repeat visits appears to be a more recent development than a time-honoured tradition. Various factors can be proposed as contributing to this apparent trend towards repetition. Perhaps the foremost is geography's shift from a spatial-descriptive exploratory enterprise to a modern scientific discipline with a premium placed on deriving generalisations through replicable results. Although few of the human geographers in the past century engaging in foreign area research have been self-conscious devotees of approaches involving nomothetic, spatial analytical science, the ascendancy of a "scientific" human geography in the post-WWII period can be implicated in the decline of chorology and geographic description for its own sake.

A second impetus also might be seen as more influence from outside geography than from within. Along with the rise of a more analytic, "scientific" human geography, geographers began to interact more with other social scientists, and borrow from their theory and practice. For field-oriented geographers working abroad, especially cultural geographers, the paths most often crossed or followed were anthropological. Increasingly archaeologists and ethnologists were digging in and making long investments in discrete study sites, rather than mounting wide-ranging surveys and broad collecting ventures. A variant of this development, especially for geographers, is repeat visit research at intervals to measure change (or lack of it) in places, landscapes, or regions.

Elements of the mobility and transience of the older survey and collection mode is carried forward, but with an emphasis on demonstrating familiarity rather than novelty. A fourth influence that began to emerge by the 1960s and was accelerated by that decade's ferment, involved ethical questions and issues. Anthropologists, followed by other social scientists including geographers, generally came to see and accept that not only their research had varying impacts on their "subjects," but also that forms of reciprocity might be part of a researcher's protocol and practice. Among other modes, this has taken the form of extended ties to individuals or communities, often spanning decades since the initial research was conducted (Richardson 1998). In varying degrees, all of these underlying factors may play a part in the decisions and the designs that have led geographers to return at intervals, often multiple times, to favoured sites, landscapes, or regions. I suspect, however, that most returnees would simply say that the main rationale or attraction is that they have come to identify strongly with their adopted site or region and feel a need or obligation to return.

In some cases it may serve as a second home, and contrary to the old saw, familiarity here breeds not contempt, but affection and attachment. James Parsons, perhaps better than anyone, evocatively makes the case for this in his AAG presidential address: "Repeated visits and growing familiarity with an area continually enlarge one's range of vision and concern. You see, you hear, or read something new, something out of place, something irregular, and you are moved to investigate further." His own periodic return visits to Antioquia in Colombia reverberated in his lifetime and continue to resonate there today. As with much of modern geography, the groundings and ground floor for both the theory and practice of foreign area fieldwork can be found in German scholarship.

By the late 19th century geographic training in German universities followed a regimen involving lecture courses, research seminars, and perhaps paramount, fieldwork. Typically a graduate student was first expected to conduct fieldwork and produce a thesis on a topic of local focus, preferably in one's Heimat or "home" locale or region. This exercise was part of the larger Heimatkunde or local study movement that was integral to the formation of German nationalism. For students completing doctoral study, a dissertation based on foreign area fieldwork was expected. In turn, foreign area studies, or Auslandskunde became an integral component of German colonial exploration and expansion as well as demonstrating the prowess of German science and scholarship to a world audience. The exemplars of the Heimatkundel Auslandskunde tradition in German geography before it became formalised in university training, were Johann Gottfried Herder and Alexander von Humboldt. Herder valorised local studies, especially of folk culture, and coined the term "nationalism" and "populism" among others.

Humboldt provided a towering example of foreign field studies with his five years of travel and investigations in South America, Mexico, and Cuba, and later in Russia and Central Asia (Livingstone 1992, Matless 1992). Both Herder and Humboldt were complex Enlightenment figures that also gave birth and voice to major tenets of Romanticism. It could be argued that their combinations of the Enlightenment penchant for empirical scientific discovery with the Romantic impulse for direct experience, especially of "nature," provided a template for future fieldwork practice in geography. One example that Humboldt did not provide, was of long-term engagement with any particular place or region. Although he could be an acute observer of local phenomena, his sights were ultimately set on framing and answering global scale questions and problems. For this, mobility and breadth of experience, not in-depth local immersion, was called for.

Moreover, he was the consummate cosmopolitan, and was content to process the wealth of empirical data that flooded in to him by his global network of colleagues from his bases in Paris and later Berlin. Herder, on the other hand, in valorising local studies, made immersion in local circumstances, including learning local languages and customs a badge of distinction. By the turn of the twentieth century, emergent national traditions in geographic practices and perspectives were beginning to coalesce in university contexts beyond Germany. The "home geography" followed by a foreign area study model may have been an ideal, but it was far from obligatory. French geographers found the geographical "other" in their own regional distinctiveness. Russian geographers could travel continental distances but still not be abroad (Bassin 1999). British geographers had the regions and resources of a global empire to construct and catalogue, but much of this could be done at home, and was. In the early years of professional geography, North American geographers largely confined their scope and ventures to national space.

There were, of course, exceptions. As early as the 1870s William Morris Davis spent three years in Cordoba, Argentina conducting research at the Argentine Meteorological Observatory. A decade later Mark Jefferson (1884-1889) also spent time at the Cordoba observatory and worked as a submanager of a sugar plantation near Tucuman before beginning his studies with Davis at Harvard (Martin 1968). Wellington Jones, co-author with Carl Sauer of an early and key statement on fieldwork (1915), assisted Bailey Willis with physiographic survey for rail lines in northern Patagonia during 1911-1912. Although these were among the first Latin American fieldwork efforts by North American geographers, none of the three made Latin America the main focus of their research and only Jefferson returned to do research.

This was in 1918 under the auspices of the American Geographical Society's ABC Expedition. Isaiah Bowman, the AGS director, commissioned Jefferson and an assistant to study European immigration in the Southern Cone, especially Germans, in the context of WWI and its aftermath. Throughout the 1920s Jefferson continued to publish the results of the several months he spent in Chile, Argentina, and Brazil in 1918, but he did not make them a return destination. Nor did he direct graduate students to the Southern Cone, as he spent his career teaching at Ypsilanti Normal School. Similarly Isaiah Bowman accumulated a wealth of experience and knowledge with his three field seasons in Andean Peru, Bolivia, and Chile.

These trips resulted in a number of important publications (1916, 1924) but like Jefferson, South America was not a repeat destination, and Bowman's role as AGS Director (1915-1935) and

President of Johns Hopkins University (1935-1949) precluded overseeing student research. Although as AGS Director he did foster and direct resources to Latin American geographic research (Martin 1980). It wasn't until academic geography's second generation in the US, or with those that began their professional careers in the post-WWI period, that sustained engagement in foreign area fieldwork took hold. For Latin America, two nodes developed coalescing around two distinctly different figures—Preston E. James and Carl O. Sauer. James came from a prosperous New England family wherein a Harvard education was not an option—it was an expectation. At Harvard he discovered geography and completed his first two degrees in geography, continuing on with doctoral studies in geography at Clark (Martin 1987). In contrast, Sauer came from a Midwestern German immigrant background, and found geography via geology in graduate school at Chicago.

Though ten years Sauer's junior, James' first forays into Latin America preceded Sauer's by several years. At the end of his first term at Clark (April 1921) he set off for an extended tour of South America. He took a United Fruit boat to Colombia, then down the Andes to Chile, through Argentina, Uruguay and Brazil, and back by boat via the Lesser Antilles. From this trip he gathered enough data and observations to publish a half dozen articles. It also gave him his dissertation topic—transportation in South America. In 1923 he replaced Sauer at Michigan, and his next field trip was to Trinidad in 1924. This resulted in another six publications. In 1930, 1938 and in 1949 as a visiting lecturer he did extensive fieldwork in Brazil. These two decades of engagement with Brazil yielded some two-dozen publications.

After WWII James travelled to a number of places in Latin America in concert with his official duties with the Pan American Institute of Geography and History (PAIGH) and in other capacities, augmenting his compendious textbook on Latin America, but like Isaiah Bowman, once he assumed a heavy load of responsibilities beyond teaching and research, fieldwork was essentially shelved, and he doesn't seem to have been concerned to return to sites once studied or observed. James' approach was broad and comprehensive, publishing on a wide range of topics from climatology and geomorphology to transportation and economic development to historical and cultural landscapes, in a number of countries. Sauer's trajectory was different. Once he left Michigan and the Midwest for Berkeley in 1923, he also developed a life-long engagement with Latin America, but only in select places and via a narrower range of geographical topics—mostly cultural and historical. Sauer's introduction to Latin America came through his musing that Baja California was "the mother of California historically" and should be studied as such.

This notion sparked interest in some of his graduate students, and Peveril Meigs and Warren Thornthwaite made a reconnaissance of the northern portion of the peninsula in the summer of 1925. Their seminar reports that fall convinced Sauer that his future lay south of the border, especially since he felt California had already been staked out and claimed by other researchers, particularly geologists, historians, and anthropologists. His first venture into Mexico was during the summer of 1926, or five years after James' first South America circuit. He went with Meigs and Fred Kniffen. They retraced Meig's itinerary the previous summer, travelling south as far as San Fernando, or about a quarter of the way down the peninsula. Sauer was primarily interested in studying the geomorphology, but Meig's interest in the Spanish missions helped turn Sauer towards the historical features in the landscape as well. Both Meigs (1932) and Kniffen (1930) went on to write dissertations on northwest Mexican topics, but neither made Latin America "their region." Sauer planned a follow-up field season in 1927 to check on Meigs and Kniffen in the field, but other things intervened.

The summer of 1928 he moved his sights east, devoting seven weeks to exploring northeastern Sonora. Plans to spend the summer of 1929 in Sonora were thwarted by political unrest there, but Sauer received a sabbatical leave for the spring of 1930 and he made the most of it. He left in late December 1929 with his family, and post-doe Gottfried Pfeifer for Mazatlan, Sinaloa. They were joined shortly by student Donald Brand and at various times during the season, anthropologists Alfred Kroeber and Paul Kirchoff and artist Antonio Sotomayor. In Sinaloa they found archaeological evidence of advanced pre-Hispanic cultures, thus Sauer was able to propose new boundaries for Mesoamerica's northwestern extension (Sauer and Brand 1932). In late March they moved their field headquarters to Nogales, Arizona. From this base they mounted "archaeogeographical" surveys of northwestern Sonora, discovering extensive trinchera (habitational terraces) settlement sites (Sauer and Brand 1931).

Sauer and Brand spent another three months in the field and also the regional archives. Winter break 1930-1931 Sauer was back in Sonora, this time accompanied by Joseph Spencer, a new graduate student from UCLA. Spencer stands out as one of Sauer's students who very much charted his own course, not following Sauer's Latin American path, but making East and Southeast Asia his career choice for fieldwork and return visits. They spent three weeks, mostly on horseback, in central Sonora on reconnaissance for trinchera sites. In early 1931, based on his previous northwest Mexican fieldwork and discoveries, Sauer was awarded a prestigious John Simon Guggenheim Memorial Fellowship to continue this work. Not only did this put Sauer back into the field for fall semester 1931, but it also led to Sauer's wider participation in the Guggenheim Fellowship programme as a member of the selection committee from 1934 to the mid-1960s. Sauer's recruitment into the Guggenheim ranks led to associations with other East Coast establishment foundations. He forged close ties with Henry Allen Moe, the Guggenheim program's secretary.

This led to introductions and involvements with the Rockefeller Foundation's Division of Social Sciences and the Carnegie Foundation, and in turn, support for Sauer and his students' Latin Americanist fieldwork. Sauer's "Guggenheim Year" (summer and fall of 1931) in Mexico expanded his experience of both field and archive to Central Mexico In the process, he became increasingly familiar with trails in the Sonora and Sinaloa. This time his graduate assistant was Leslie Hewes, who like Spencer, did not elect Latin

America for career commitment, choosing instead a Heimat theme—the Cherokee country in his own native Oklahoma for his dissertation and then the wider Great Plains for his subsequent work (Wishart 2000). Sauer was not successful in finding the sought-after evidence for demonstrating links between the high culture areas of western Mexico and the US Southwest, one of the objectives of his Guggenheim fellowship. He did, however, gather material for his acclaimed study, "The road to Cibola" (1934b), an interpretation of sixteenth-century Spanish exploration and travel along the west coast of Mexico northward to the US Southwest.

This publication, more than his previous efforts, brought Sauer to the attention of historians and others with interests in colonial exploration. His next venture took him to Chihuahua and Durango in the summer of 1933. Along with a graduate assistant, Sauer covered a lot of ground, visiting Mennonite as well as Tarahumara settlements, and "discovering" the Parral archives. The archival material from Parral and other local archives yielded two important papers one on aboriginal tribes and languages and the other on aboriginal population (Sauer 1934a, 1935). In effect, the population study was the founding document in the formation of the other "Berkeley school"—that of historical demography. His turn in 1934 to new commitments "back East," which included not only the boardrooms of East Coast foundations but also the corridors of New Deal Washington, especially the new Soil Erosion Service, diverted his Mexican attentions. Still, he managed to spend the summer of 1935 in the field, this time exploring the western edge of the central Mexican plateau, and spending a month in the archives of Mexico City.

Robert Bowman, Isaiah Bowman's son and new Berkeley graduate student joined him in the field. As with Spencer and Hewes, Bowman's summer with Sauer didn't direct him to a follow up season in Mexico, though he did do a study of soil erosion in Puerto Rico for his dissertation (Bowman 1941). Sauer's eastern involvements and administrative duties on campus kept him out of the field for the next three years. By Christmas break 1938-1939 he was back into the field, returning to Sinaloa to visit on-going archaeological excavations by Gordon Ekholm and by Isabel Kelly. That summer he concentrated his attention on Colima, investigating human-plant interactions, the archaeogeography of the region, and historical materials. The results are recorded in his historical monograph, Colima of New Spain in the Sixteenth Century (1948). Henry Bruman, one of his students working on aboriginal beverages, joined him for part of the Colima season.

The accumulated wear and tear from more than a decade of annual trips to Mexico caught up with Sauer after the summer of 1939. Much of the 1939—1940 academic year Sauer was taken up recovering from amoebic dysentery and other undiagnosed maladies.

By early 1941 he was well enough to take another sabbatical leave, and once again it was in Mexico. This time the focus was on plant use and craft production in west-central Mexico, as well as a month spent in the Mexico City archives. No direct publications resulted from this season, but Sauer's (1941) seasoned essay "The personality of Mexico" appeared during this time. It has become something of a testament to the benefits and rewards of getting to know a place, region, or nation through periodic immersion. And his growing interest in plant-human relations, observed first-hand in Mexico, provided a base for his seminal Agricultural Origins and Dispersals (1952). The other base for this book was his Rockefeller Foundation sponsored six-month tour of western South America in 1942 (West 1982).

As his research assistant, he took his son Jonathan, at the time a geography graduate student at the University of Wisconsin. Save for an Office of Naval Research-related trip in 1952 to the Caribbean, this South American transect was Sauer's only deflection from his multi-year devotion to fieldwork in many regions of Mexico. The war years 1942-1944 kept Sauer desk bound more than was his custom, but he did return to Mexico at the end of 1944 on a Rockefeller grant, and during the winter and spring of 1945 he was back to familiar locales as well as new ones. He spent a month in Oaxaca and Puebla, parts of Mexico he hadn't previously visited. Following WWII Sauer went back to his annual trips to

Mexico. Just as he had pushed farther south into Latin America in increments over two decades, he had recalibrated his estimates of the time depths of various cultural historical processes. He had become convinced that human occupancy of the Americas was far earlier that consensus opinion allowed, and that Mexico's arid landscapes might yield new evidence.

With a 1946 spring leave from teaching, Sauer and the archaeologist Emil Haury, made a transect of Sonora looking for signs of early humans. That summer he went on a similar search down the Baja Peninsula with graduate students William Massey (archaeology) and Edwin Hammond (physical geography). Sites were found that Sauer interpreted as encouraging, and returned in early 1947 with an interdisciplinary group of Berkeley faculty and students under the auspices of the newly organised "Associates of Tropical Biogeography." Sauer's agenda was not necessarily shared, and the faculty geologist, botanist, and paleontologist all preferred to pursue their own interests. That summer he took the first of several trips in which Mexico was traversed from north to south, looking for early human sites, but also of course, other things that caught his eye.

The group included Berkeley colleagues Sherbourne Cook and Leslie Simpson, Jonathan Sauer (now working with Edgar Anderson at Washington University), and two young graduate students, Jesse Walker and Rees Walker (not related). In the summer of 1948 Sauer performed a repeat of the previous year's transect, and again made it as far south as Chiapas, but turned back due to road conditions. Along on the trip were James Parsons, recent Berkeley PhD and graduate student David Lowenthal. Sauer's last trip to Baja California was during another sabbatical leave, in the winter of 1949. This time he eschewed faculty colleagues and took three graduate students, Brigham Arnold, Homer Aschmann, and Thomas Pagenhart. Arnold (1954) settled on a dissertation topic that carried forward Sauer's pursuit of early human occupancy, while Aschmann (1954) wrote his dissertation on the historical demography and ecology of Central Baja. Pagenhart demurred, filing a dissertation two decades later on historical California water issues. Sauer's final "grand tour" in the summer of 1949 was with graduate students Douglas Powell and Clyde Patton.

Among other objectives were to drive on to Guatemala, but the road ended at the border, and rail transport of the truck was not possible. His final field excursion to Mexico was the summer of 1950. The objective was to study the Pleistocene history of Parras Basin, Coahuila, searching for early human sites, and to find dissertation topics for graduate students Philip Wagner and John Vann. Once they were in place working on the vegetation history and geomorphology and of the basin, Sauer joined the pine expert Nicholas Mirov and Pagenhart on their way south to study pine forests. Among these latter students, only Vann (1959) went on to develop a dissertation out of the trip. Vann was also the last of Sauer's students to file a dissertation on Mexico, save Leonard Sawatzky (1967), who studied Mennonite colonisation, and filed his a decade after Sauer's retirement.

Among Sauer's students who made Mexico a regular destination, Robert West, Donald Brand, and Campbell Pennington stand out. West's dissertation (1948) was on the Parral mining district, but he had "discovered" the archives independently while conducting fieldwork for his master's thesis in geography at UCLA. As he remarked to me (personal communication, 25 May, 1989), Sauer had not so subtly tried to dissuade him from "mining" these archives (presumably so Sauer could do it himself), but West stuck to his guns, and produced a model study. West made over thirty field trips to Mexico from 1936 to 1990, or a span of more than fifty years. His last book on Mexico, Sonora, Its Geographical Personality (1993), distilled a half-century of observations of people and place. Brand (1951; 1966) also made Mexico his primary locus and focus, from 1930 until after retirement in late 1970s, or almost a half-century.

Pennington's dissertation (1959) was directed by James Parsons, but he was very much a Sauer protege. His field excursions in northern Mexico starting in the 1950s and continuing through the 1970s, often on mule back to record the facts of indigenous language and landscapes, were legendary. They yielded a quartet of books

(1963; 1969; 1979; 1980) that are monuments to intensive localised fieldwork. A second group that must be considered in the dynamics of repeat visitation research, aside from the repeat researcher and his or her students that also engage the site, are the researchers inspired by the example, but who may or may not be direct legatees. As such, they might be seen as repeat visit proxies. Perhaps the two best examples are Isabel Kelly and William Doolittle. As a Berkeley anthropology graduate student, Kelly took classes with Sauer, but he was not her advisor.

Doolittle is a fourth generation Sauer descendent (Brown and Mathewson 1999). For both, Mexico became their principle focus of fieldwork and research. Isabel Kelly was the second woman to receive a PhD (1932) in anthropology at Berkeley. Sauer had become her unofficial mentor in graduate school, and afterwards Kelly became a long-standing figure in the shifting retinue of participants in his Mexican fieldwork. He helped secure funding for her archaeological excavations in western Mexico, and in turn she focused on questions that furthered Sauer's cultural-historical programme, whether cultural boundaries, plant domestication and use, or early human occupations. In this sense, she represents a particular kind of proxy—one that advances a researcher's programme while remaining in situ over an extended period—in this case over several decades devoted to western Mexican archaeology and culture history.

The other kind of proxy is much more common. This is the researcher, whether directly connected to the initiator or not, who continues a line of research begun by someone else in a particular place or region. In effect, he or she performs repeat visitation geography as a proxy for the principle, though the objectives and methods of the research may shift over time. Much of Doolittle's (1988; 1990) research in Mexico has refined and expanded the programme of archaeogeography that Sauer and Brand initiated in northwest Mexico in the 1930s. His record of repeat visitations is unequalled, and not likely to be surpassed in the future, as it is still ongoing. He has made more than sixty research trips to Mexico, starting in 1977 and spanning more than thirty years. He has made seventeen trips to Sonora, or about one every other year over this thirty-year period.

What Sauer and Brand began in Baja California in the late 1920s, is being continued by Doolittle as well as other geographers. This kind of reiterative practice can produce minor research traditions resistant to change, but it can also allow for multi-decadal engagements by initial researchers or their proxies—that provide windows on site and situational change. Perhaps one of the most encouraging developments in Latin Americanist geography from the perspective of committed "returnees," is the interest being shown by a new generation of Latin American geographers, especially in Mexico and Colombia, in the work of Sauer and his students. Berkeley school classics are being translated, disseminated, and one assumes, read. It is probably too early to tell, but chances are that a new cadre of proxies is being formed in this case, researchers who are not distant "returnees," but ones who culturally at least, have always already been half way or more there.

TECHNIQUES FOR MODERN LARGE-SCALE GEOMORPHIC ANALYSIS

It can be argued that the methodologies of a science are reflected in the research techniques of its practitioners. Indeed the recent review by Goudie clearly demonstrates the methodological focus of geomorphology on small-scale shortduration process studies. However, proper techniques do not guarantee proper results. Büdel provided a parable of a misguided process geomorphologist "of a generation who no longer read A. Penck."

This fictitious scientist studied processes on Alpine upland surfaces by "modern" methods, including soil analysis, grain-size distributions, clay mineralogy, slopewash monitoring, morphometry, and statistical analysis. Büdel observes: "His conclusion was that these processes created the trough shoulders of the Alps.

His evidence for the certitude of these results was the indubitable precision of the analysis." The neglected fact was that the measured modern processes are all completely ineffective in modifying landforms that are relict from ancient times and that were formed by processes controlled by a completely different climate from that prevailing today.

Role of Space Technology

Modern macrogeomorphology makes extensive use of global observations from spacecraft that employ a variety of imaging and sensing systems. These include vidicon imaging, multispectral scanning, radiometers, and radars. Modern image processing of digitally formatted data has revolutionized the interpretation of large-scale planetary landscape scenes. The following question is posed: why have geomorphologists been slow to appreciate the global perspective afforded their planet by these advances? The question probably has many answers. The technology of remote sensing has only recently advanced to the point at which many geomorphologists can appreciate its relevance. The technology requires training in disciplines not normally considered in the training of geomorphologists. Even more interesting is the requirement placed on the user to have a large-scale view of problems.

Consider the perspective of a fluvial geomorphologist interested in floods. One approach might be to measure in excruciating detail the flood events he can easily access from his temperate climate university. This provides an impressive data set, but a problem remains. What of the immense rare floods that affect the great tropical regions of the planet? The immensity of flood effects in southwestern Queensland that followed phenomenal rains in 1974. Without satellite images, this scale of flooding could not have been measured at all.

Another solid example of the utility of appropriately processed space imagery is strikingly portrayed. This Landsat image of block mountains and basin fill in the Mojave Desert on the California and Arizona sides of the Colorado River between Lake Mohave and an agricultural development south of Parker. It has received both computer enhancement and special photographic reproduction. The resulting colour composite brings out both pronounced and subtle colour tones and patterns in the alluvial fill and pediments. In some parts of the scene, individual wash or fan patterns can be traced back to immediate source rock areas.

Three broad colour classes of unconsolidated surface materials can be recognized:

1. Light buff to tan material commonly in lower levels of the basin,
2. Reddishbrown material, and
3. Dark gray to blue gray material, forming aprons around many eroding ranges and generally overlapping the first class.

Although this false-colour rendition departs from the normal colours observed in the field or in natural colour aerial photography, it can be interpreted in terms of parent lithologies. Both the colour emphasis and the interactions among alluvial of deposits emanating from the mountain uplifts are far better displayed and interpreted synoptically than has been possible from aerial photographs, even after these are joined in mosaics.

In any science, new techniques are not important of themselves. It is rather the new discoveries made possible because of those techniques that stimulate scientific progress. A profound example of such a new discovery in terrestrial geomorphology came in November 1981 when the shuttle Columbia trained a space-age instrument on the Earth.

The Shuttle Imaging Radar carried by Columbia produced radar images of the hyperarid Selima Sand Sheet on the eastern Sahara.

The radar penetrated the sand cover to reveal fluvial valleys now filled by eolian sand. The valleys discovered by radar interpretation show a regional drainage system formed when the modern eolian-dominated landscape was subject to extensive fluvial erosion, probably during pluvial episodes of the Pleistocene and Tertiary.

DYNAMICS OF GEOMORPHOLOGY

To place geomorphology upon sound foundations for quantitative research into fundamental principles, it is proposed that geomorphic processes be treated as gravitational or molecular shear stresses acting upon elastic, plastic, or fluid earth materials to produce the characteristic varieties of strain, or failure, that constitute weathering, erosion, transportation and deposition.

Shear stresses affecting earth materials are here divided into two major categories: gravitational and molecular. Gravitational stresses activate all downslope movements of matter, hence include all mass movements, all fluvial and glacial processes. Indirect gravitational stresses activate wave-and tide-induced currents and winds. Phenomena of gravitational shear stresses are subdivided according to behavior of rock, soil, ice, water, and air as elastic or plastic solids and viscous fluids. The order of classification is generally that of decreasing internal resistance to shear and, secondarily, of laminar to turbulent flow.

Molecular stresses are those induced by temperature changes, crystallization and melting, absorption and desiccation, orosmosis. These stresses act in random or unrelated directions with respect to gravity. Surficial creep results from combination of gravitational and molecular stresses on a slope. Chemical processes of solution and acid reaction are considered separately.

A fully dynamic approach requires analysis of geomorphic processes in terms of clearly defined open systems which tend to achieve steady states of operation and are self-regulatory to a large degree. Formulation of mathematical models, both by rational deduction and empirical analysis of observational data, to relate energy, mass, and time is the ultimate goal of the dynamic approach.

Geomorphology

The surface of Earth is modified by a combination of surface processes that sculpt landscapes and geologic processes that cause tectonic uplift and subsidence. Surface processes comprise the action of water, wind, ice, fire, and living things on the surface of the Earth, along with chemical reactions that form soils and alter material properties, the stability and rate of change of topography under the force of gravity, and other factors, such as (in the very recent past) human alteration of the landscape. Many of these factors are strongly mediated by climate. Geologic forcings include the uplift of mountain ranges, the growth of volcanoes, isostatic changes in land surface elevation (sometimes in response to surface processes), and the formation of deep sedimentary basins where the surface of Earth drops and is filled with material eroded from other parts of the landscape. The Earth surface and its topography therefore are an intersection of climatic, hydrologic, and biologic action with geologic processes.

The broad-scale topographies of Earth illustrate this intersection of surface and subsurface action. Mountain belts are uplifted due to geologic processes. Denudation of these high uplifted regions produces sediment that is transported and deposited elsewhere within the landscape or off the coast. On progressively smaller scales, similar ideas apply, where individual landforms evolve in response to the balance of additive processes (uplift and deposition) and subtractive processes (subsidence and erosion).

Often, these processes directly affect each other: ice sheets, water, and sediment are all loads that change topography through flexural isostasy. Topography can modify the local climate, for example through orographic precipitation, which in turn modifies the topography by changing the hydrologic regime in which it evolves. Many geomorphologists are particularly interested in the potential for feedbacks between climate and tectonics mediated by geomorphic processes.

In addition to these broad-scale questions, geomorphologists address issues that are more specific and/or more local. Glacial geomorphologists investigate glacial deposits such as moraines, eskers, and proglacial lakes, as well as glacial erosional features, to build chronologies of both small glaciers and large ice sheets and understand their motions and effects upon the landscape. Fluvial geomorphologists focus on rivers, how they transport sediment, migrate across the landscape, cut into bedrock, respond to environmental and tectonic changes, and interact with humans.

Soils geomorphologists investigate soil profiles and chemistry to learn about the history of a particular landscape and understand how climate, biota, and rock interact. Other geomorphologists study how hillslopes form and change. Still others investigate the relationships between ecology and geomorphology. Because geomorphology is defined to comprise everything related to the surface of Earth and its modification, it is a broad field with many facets.

Practical applications of geomorphology include hazard assessment (such as landslide prediction and mitigation), river control and stream restoration, and coastal protection.

History

With some notable exceptions, geomorphology is a relatively young science, growing along with interest in other aspects of the Earth Sciences in the mid 19th century. This section provides a very brief outline of some of the major figures and events in its development.

Ancient Geomorphology

Perhaps the earliest one to devise a theory of geomorphology was the polymath Chinese scientist and statesman Shen Kuo (1031-1095 AD). This was based on his observation of marine fossil shells in a geological stratum of a mountain hundreds of miles from the Pacific Ocean. Noticing bivalve shells running in a horizontal span along the cut section of a cliffside, he theorized that the cliff was once the pre-historic location of a seashore that had shifted hundreds of miles over the centuries.

He inferred that the land was reshaped and formed by soil erosion of the mountains and by deposition of silt, after observing strange natural erosions of the Taihang Mountains and the Yandang Mountain near Wenzhou. Furthermore, he promoted the theory of gradual climate change over centuries of time once ancient petrified bamboos were found to be preserved underground in the dry, northern climate zone of *Yanzhou*, which is now modern day Yan'an, Shaanxi province.

INTRODUCTION TO THE FIELD SITE

As a single lab exercise, before we go to the specific field site, A twofoot-contour interval topographic map of an area of Toby Creek and indicate the field area where they will develop their chronosequence. In this portion of Toby Creek, there are several relatively flat concordant upland surfaces that students map. We review elements of topographic maps, and students convert the scale of the map from feet to metres (an exercise that seemingly cannot be repeated enough at this academic level).

The exercise ultimately requires students to develop a geomorphology map and a topographic profile of a transect across the field area. As a class, we qualitatively note areas of similar morphology (flat, vs. steep vs. very steep) and make hypotheses about origin of the landforms that the map is depicting (floodplain, possible terraces, tributary alluvial fans, and anthropogenically modified areas). The concept of 'terrace' and 'alluvial fan' loosely here as the scale is rather small, and we emphasise that by examining the map, we are developing hypotheses of stratigraphy and landform origin that can be tested in the field.

The concept of a mapping unit and explain how a geomorphologie map differs from a geologic map. Together we define mapping units for the field area, and students then individually colour in their maps using provided possible mapping units and make a key to their units. Students are graded primarily on the accuracy of their topographic profiles and their ability to follow the provided 'rules' of geologic mapping. For example, all polygons must be closed or go off the map, all areas of the map must be assigned a mapping unit, and all maps must be neat with consistent and correct labelling.

2005 Project Example

In 2005, students mapped two tributary alluvial fan surfaces and two 'upland surfaces' in their designated field area. Field relationships indicated that one alluvial fan was inset into another so a relative age relationship was established, with Alluvial Fan 1 hypothesised to be older than Alluvial Fan 2. It was hypothesised that the upland surfaces were related to long-term fluvial incision, and thus the higher of the two surfaces (Upland Surface 1) was hypothesised to be older. For logistical reasons, we chose to focus our study on the tributary alluvial fans and the fioodplain. Five pits were excavated and described by each group of students. Pit #1 was located on the fioodplain. Pits #2 & #3 were located on Alluvial Fan 2, whose surface is approximately two metres above that of the fioodplain. Pits #4 & #5 were located on Alluvial Fan 1, whose surface sits approximately 4 metres above the fioodplain.

These are typical of any year's class in that they range from purely illustrative to quantitative.. The sophistication of the visuals also is variable, but hand-drawn graphs are accepted and allow less-computer savvy students to feel on equal footing. The following is a combined summary of the results presented by students for the 2005 project. The detail of the results is illustrative of the complexity of the ideas and conclusions that students are able to make using the data that they have collected. With only minimal change in emphasis of data collection and analysis, the same project could have highlighted other geologic disciplines such as geomorphology, sedimentology or mineralogy.

For the 2005 project, there were notable differences between soils on different surfaces and between soils on the same surface. The fioodplain deposits expectedly exhibited relatively weak overall soil development with A-BwC horizonation with a buried soil at depth overlying undisturbed fluvial stratigraphy.

Students noted that the surface soil in this pit exhibited slightly redder colours and more structure than the cutbank exposure they had previously examined. Also, the soil in Pit 1 was lacking the unweathered sand deposit from a 2004 flood which we had also observed in the cut-bank exposure. These observations led students to examine the location of the pit in more detail and to recognise that the area in question was potentially bypassed from flooding due to the presence of a crevasse splay in the nearby levee. I used this and other similar experiences to emphasise to students that we can often learn more when a hypothesis that we propose (in this case, soils developing on a single geomorphic surface should exhibit similar characteristics) is proved false.

The soils exposed by Pits 2 & 3 on the lower alluvial fan exhibited similar characteristics. Both soils are developing in homogeneous, well-sorted grussy-sand that is consistent with a tributary alluvial-fan origin for this surface. The soils exhibited diffuse relatively dark A horizons which extend to approximately 20 cm to 40 cm depth. These two soils also exhibit approximately 20-40 cm thick oxidised B-horizons with few if any clay films evident. Colours at the bottom of these pits were approaching the fioodplain sediment.

Pits 4 and 5 shared both similarities and differences. Pit 4 and the upper horizons of Pit 5 were characterised by similar well-sorted sand as pits 2&3. The A horizons for these pits were much thinner (only 6-10 cm in depth) than those of Pits 2& 3 and were also much more concentrated in organic material as evidenced by darker colours. These A horizons overlie a discrete Bt horizon of 10-15 cm thickness with welldeveloped clay films developing in the sandy parent materials. In Pit 4, this Bt horizon graded into a well-oxidised B horizon with occasional clay films present. Underlying a relatively thin Bt horizon in Pit 5, however, was extremely weathered, clay-rich saprolite with extensive mottling, gleying, and clay films. The overall soil development of this portion of Pit 5 was significantly greater than the other four soils examined. This difference in soil development is reflected in a high overall Profile Development Index.

In the case of the 2005 study, most groups recognised the overall increase in soil development in deposits of increasingly higher elevation. There were notable exceptions to this trend, however, and these were discussed by many groups in their presentations. For example, many students noted the presence of Mn nodules in all pits regardless of age. Thus, students were able to conclude that the presence of Mn is not a good indicator of soil age.

The group that calculated clay film indices noted an unpredicted relative decrease in clay films from the floodplain soil to the soils on the lower alluvial fan. It was suggested that the decrease was due to the fact that the buried soil in Pit 1 was used as a parent material reference, likely not an appropriate choice for their index. All groups noted that Pit 5 was the only pit prone to filling up with water after a rainfall. Many groups concluded correctly that the high clay content of this pit is contributing to periods of sustained saturation, and thus red-ox conditions, as evidenced by the extensive mottling observed there.

Some groups noted the differences in A horizon morphology between Pits 2&3 and Pits 4&5 and after some prompting were able to draw some conclusions. The relatively high permeability of the sandy matrix in Pits 2&3 permitted dark organic material to be transported to deeper portions of the soil profile, thus resulting in relatively diffuse, thick A horizons. It was suggested that this transportation was not possible in the soils of Pits 4&5 because sufficient clay had accumulated to retard infiltration into these soils. Thus the soil development of the B horizon resulted in a change in the development of the A horizon. A final interesting conclusion from the 2005 project was the recognition that deposition of tributary alluvial fans in the Toby Creek drainage is episodic, as is evidenced by fans with different degrees of soil development. Pit 5 provided evidence that the thickness of these fans is variable and that an older, more weathered landscape is preserved beneath them.

FIELD EXPERIMENTS IN GEOMORPHOLOGY

Fieldwork is one of the central methods of investigation through which human geographers gather information about people, places, and landscapes, and generate formal knowledge about space–society relations in different contexts. Many human geographers identify themselves and the evolving contexts of their work through fieldwork practices. Quite often, fieldwork is one of the research activities that other academics and the general public outside the discipline most associate with human geography. Fieldwork has long been considered a rite of passage among human geographers, though there has been disagreement regarding the importance of fieldwork to the identity of the discipline. While fieldwork practices, such as analyses of landscapes, have been common in regional and cultural geography and considered by many to be among the building blocks of the discipline, other fieldwork approaches such as ethnography and participant observation have not been central to human geography until more recently. The disagreements voiced by some geographers regarding the role of fieldwork in geographic research seem to stem from the variety of practices that geographers count as 'fieldwork'. In spite of these different assessments, the majority of the systematic branches of the discipline today, from economic and urban geography to cultural and critical geography, consider fieldwork as an important component of research endeavors in geography.

Field experiences are a critical part of a geoscience student's education and have long been one of the foundations of geoscience education. Below, you will find links to presentations, posters, and teaching activities presented by participants at several workshops that speak to a variety of ways they have taught geomorphology with and through field work.

3

The Geomorphic System

Geomorphologists will develop better understanding of the scientific foundations of their field through philosophical analysis. Although philosophical self-examination could be conducted in an independent manner, it is best pursued in the context of ongoing debate in the philosophy of the physical sciences. This type of analysis allows geomorphologists to address questions such as: what constitutes a scientifically valid explanation; are there different types of acceptable explanations, and, if so, what relationships, if any, exist among them; what is a geomorphological theory; to what extent is geomorphological knowledge a product of social, cultural, political, and ethical factors; are geomorphologists problem solvers or truth seekers; how do geomorphologists justify and defend their claims to knowledge; are these methods consistent with the aims of the field; and how are responses to these questions similar to or different from those provided by other scientific disciplines?

Although social constructivism, postpositivist empiricism, and scientific realism are the primary focuses of current philosophical debate, other viewpoints also may contain ideas relevant to geomorphology. Specifically, it may be worthwhile to explore philosophical perspectives that have emerged in disciplines such as biology and physics. However, there is no reason to presuppose that a philosophical framework for geomorphology will be merely a restatement of the philosophy of another discipline. Because geomorphology is concerned with distinctive types of natural systems that include synergistic physical and biological elements and employs characteristic investigative methods, it cannot be reduced to the underpinning disciplines.

One predominant idea in the philosophy of science merits serious consideration by geomorphologists. The three contemporary philosophical perspectives discussed above, though widely divergent in many respects, uniformly subscribe to the theory-laden view of scientific observation. Given their many differences, this commonality suggests something of uncommon significance. It also contrasts sharply with the theory-neutral interpretation of observation that appears to be a de facto—albeit often unconscious—belief of many geomorphologists. Philosophers have devoted considerable attention to the topic of scientific observation. Although a universally accepted definition has not emerged from this analysis, most philosophers agree that scientific observation cannot be equated only with perception; instead, it is a complex process that includes identification, interpretation, and description. The relationship between theory and observation in geomorphology deserves more consideration than it has received.

The philosophy of science not only provides a means of evaluating contemporary geomorphology but also serves as a framework for assessing the historical development of the discipline. Most available philosophical perspectives incorporate the explicit caveat that they apply to mature sciences only. Even scientific realists acknowledge that scientific theories may be largely social constructs in an inchoate science; therefore, it becomes difficult to assess whether a scientific discipline corresponds to the realist conception of scientific progress until it has developed a substantial history. Whether or not geomorphology qualifies as a mature science is unclear; nevertheless, any tentative assessment or debate about the nature of such an assessment should occur in the context of existing philosophical ideas.

In geomorphology the transition from the Davisian view of landscape development to the modern emphasis on systematic process-oriented investigations has been explicitly and implicitly characterised as a Kuhnian-style conceptual change or paradigm shift. On the assumption that the majority of geomorphologists prefer to view their field from a realist perspective, this characterisation of the historical development of geomorphology is inconsistent with the prevailing philosophical perspective, because the Kuhnian theory of science incorporates the notion that science does not progress towards the truth. The challenge for realists is to show how many theoretical constructs embodied in the Davisian view of geomorphology, including references to unobservables, have been preserved in contemporary geomorphic theories (Rhoads forthcoming). On the other hand, process geomorphologists must recognise that adopting a realist perspective does not necessitate that a truly scientific approach implies that all geomorphological problems must be described in the language of physics. Contemporary scientific realism explicitly acknowledges that no scientific discipline has privileged status with regard to the truth.

Exploration of philosophical issues in geomorphology should also enhance disciplinary unity. Contention in contemporary geomorphology centres on differences in regulative principles, types of scientific arguments, and characteristics of theory employed by scientists, all of whom consider themselves geomorphologists (Rhoads and Thorn 1993). Although the point of contention is scientific methodology, the contention itself is clearly philosophical in nature. In other words, it is not possible to resolve conflicts between competing methodologies within science itself; instead, resolution of these differences must occur within philosophy (Montgomery 1991). As this issue and others like it are examined philosophically, differences in opinion will emerge about which viewpoint provides the clearest perspective on geomorphology. However, here the goals of science and those of philosophy differ. Whereas practicing scientists define the success of their endeavours in empirical terms, such as confirmation and acceptance of factual knowledge, philosophers deal with concepts and define success quite differently.

Philosophers commonly exhibit great respect for the intellectual contributions of people who hold ideas directly opposed to their own. Success in philosophy is measured not by declaring that one perspective is right and that others are wrong, as tends to happen when methodological issues are debated in science, but by the degree to which opposing viewpoints have reached a consensus as to which issues are central to a specific debate and by the level of rational sophistication of the arguments.

Philosophical insight by necessity depends on intellectual disagreement and debate. Viewed in this light, philosophical introspection provides an excellent antidote to scientists who wish to divide their colleagues into winners and losers on the basis of methodological preferences. Indeed, it offers the opportunity to embrace methodological diversity in a substantive and constructive manner and to enhance the intellectual depth of the discipline.

MORPHOLOGIC AND CASCADING SYSTEM

Morphological systems: The network of structural relationships or cross-correlations between the constituent parts of systems.

Cascading systems: The path followed by throughputs of energy or mass.

Morphological System - this is a system where we understand the relationships between elements and their attributes in a vague sense based only on measured features or correlations. In other words, we understand the form or morphology a system has based on the connections between its elements. We do not understand exactly how the processes work to transfer energy and/or matter through the connections between the elements.

Cascading System - this is a system where we are primarily interested in the flow of energy and/or matter from one element to another and understand the processes that cause this movement. In a cascading system, we do not fully understand quantitative relationships that exist between elements related to the transfer of energy and/or matter.

Characteristics of systems in Geography

Most systems in Geography share the same common characteristics. These common characteristics include the following:

1. Systems are a generalisation of reality – they give us an idea of what is happening in the system but the reality is often more complex and requires detailed study
2. They have a structure that is defined by its parts and processes.
3. Systems tend to function in the same way. This involves the inputs and outputs of material (energy and/or matter) that is then processed causing it to change in some way.
4. The various parts of a system have relationships between each other. They are often Connected and integrated together.
5. The fact that functional relationships exist between the parts suggests the flow and transfer of some type of energy and/or matter.
6. SOME systems often exchange energy and/or matter beyond their defined boundary with the outside environment, and other systems, through various input and output processes.
7. Functional relationships can only occur because of the presence of a driving force.

Cascading System

Cascading system In geomorphology, a type of dynamic system characterized by the transfer of mass and energy along a chain of component subsystems, such that the output from one subsystem becomes the input for the adjacent subsystem. An example is the valley glacier, where the inputs of snowfall and rock debris from the slopes above, and potential energy (derived from elevation) are cascaded through a sequence of climatic environments with a progressive reduction in mass and dissipation of energy, the output from the glacier being sediment and water which form the input to the proglacial subsystem.

GENERAL SYSTEM THEORY

Geomorphic Systems is the study of deep and shallow Earth processes that integrate through time to shape the landforms and landscapes that compose our physical environment. Once the link between process and landscape is understood, then we can read the landscape to interpret the present and past Earth processes active in a region.

The societal applications for that knowledge include land-use planning, geologic hazard mapping, ecosystem restoration and predicting the effects of global climate change. Ecosystem restoration includes either reconstructing an equilibrium landscape in a disturbed site, or encouraging the surface processes that will form the equilibrium landscape over time.

Therefore, the practice of ecosystem restoration requires a fundamental understanding of the intimate links between earth processes and the landforms they construct. Global change affects rates and styles of geomorphic change, therefore, we can read paleoclimate from the soils and landforms we study.

Systems of all kinds are open, closed, or isolated according to how they interact, or do not interact, with their surroundings. Traditionally, an isolated system is a system that is completely cut off from its surroundings and that cannot therefore import or export matter or energy.

A closed system has boundaries open to the passage of energy but not of matter. An open system has boundaries across which energy and materials may move.

All geomorphic systems, including hill slopes, are open systems as they exchange energy and matter with their surroundings. They are also dissipative systems, which mean that irreversible processes resulting in the dissipation of energy (generally in the form of friction or turbulence) govern them. Thus, to maintain itself, a geomorphic system dissipates energy from such external sources as solar energy, tectonic uplift, and precipitation. Four kinds of geomorphic system may be identified: form systems, process systems, form and process systems, and control systems.

Process systems, which are also called cascading or flow systems are defined as 'interconnected pathways of transport of energy or matter or both, together with such storages of energy and matter as may be required'.

An example is a hill slope represented as a store of materials: weathering of bedrock and wind deposition add materials to the store, and erosion by wind and fluvial erosion at the slope base removes materials from the store. The materials pass through the system and in doing so link the morphological components.

4

Concepts of Geomorphology

Most geomorphologists would agree that certain fundamental assumptions underlie all geomorphological investigations. Whether termed "fundamental concepts", "philosophical assumptions", "paradigms", or "basic postulates", these ideas constitute a "conventional wisdom" for the science. One such fundamental concept involves the inherent complexity of landscapes. This concept has impeded the development of grand theories that survive the test of explaining numerous local features.

Another basic assumption involves climatic morphogenesis, emphasizing the role of climatically controlled processes of landform genesis. Several of these concepts have yielded major intellectual controversy, such as the role of cataclysmic processes in shaping the landscape. These concepts apply to geomorphology of all scales.

Geomorphologists are also "landscape-detectives" working out the history of a landscape. Most environments, such as Britain and Ireland, have in the past been glaciated on numerous occasions, tens and hundreds of thousands of years ago. These glaciations have left their mark on the landscape, such as the steep-sided valleys in the Lake District and the drumlin fields of central Ireland. Geomorphologists can piece together the history of such places by studying the remaining landforms and the sediments – often the particles and the organic material, such as pollen, beetles, diatoms and macrofossils preserved in lake sediments and peat, can provide evidence on past climate change and processes.

So geomorphology is a diverse discipline. Although the basic geomorphological principles can be applied to all environments, geomorphologists tend to specialise in one or two areas, such aeolian (desert) geomorphology, glacial and periglacial geomorphology, volcanic and tectonic geomorphology, and even planetary geomorphology. Most research is multi-disciplinary, combining the knowledge and perspectives from two contrasting disciplines, combining with subjects as diverse as ecology, geology, civil engineering, hydrology and soil science.

The Nature Of Geomorphology

Geomorphology is the study of landforms and landscapes, including the description, classifiation, origin, development, and history of planetar surfaces. During the early part of this century, the study of regional-scale geomorphology was termed "physiography". Unfortunately, physiography also became synonymouswith physical geography, and the concept became embroiled in controversy surrounding the appropriate concerns of that discipline. Some geomorphologists held to a geological basis for physiography and emphasized a concept of physiographic regions, A conflicting trend among geographers was to equate physiography with "pure morphology," divorced of its geological heritage.

In the period following World War II, the emergence of process, climatic, and quantitative studies led to a preference by many Earth scientists for the term "geomorphology" in order to suggest an analytical approach to landscapes rather than a descriptive one. In the second half of the twentieth century, the study of regional-scale geomorphology—the original physiography-was generally neglected. Russell attributed the decline of physiography to its elaborate terminology and to its detachment from evidence acquired by other disciplines, chiefly geology.

Although the concept of physiographic regions endured among geologists, geographers became much more interested in the details of man/land interactions and in the applications of modeling and systems analysis to geomorphology. In the exploration of planetary surfaces by various space missions, the perspective of regional geomorphology has been the required starting point for scientific inquiry. Global studies of Mars, the Moon, Mercury and Venus resulted in the identification of "surface units" or physiographic provinces.

The Colorado Plateau is an excellent example of a terrestrial physiographic province. Plate I-1 illustrates the use of a large-scale perspective to focus on this naturally defined region. The term "mega-geomorphology" was introduced in March 1981 at the 21st anniversary meeting of the British Geomorpholgy Research Group. The proceedings of that meeting reveal that the concept was not well defined. It clearly involves a return by geomorphologists to the study of phenomena on large spatial scales, ranging from regions to continents to planets.

It also involves large time scales. Nevertheless, mega-geomorphology is merely a convenient term, unencumbered by past philosophical trappings, that emphasizes planetary surface studies at large scales. The interrelation of temporal and spatial scales in geomorphology is illustrated by the tentative classification. Of course, such a hierarchial ordering of geomorphic features is far from satisfying. As stated by Sparks, classifications are arbitrary constructions designed to facilitate the discussionof diverse phenomena at the risk of some distortion of the truth. The scheme merely illustrates what was well known to the great geomorphologists at the last turn of the century. The large first order features, continents and ocean basins, persist and evolve over long time scales. Small high-order features are transient. Fundamental units appear at different orders.

The old concept of physiographic regions was used to designate second-order forms, such as entire mountain ranges or coastal plains. Massive entities within a physiographic region might constitute a third-order form, such as a domal uplift. The details of the classification are unimportant as the analysis moves on to exploring the explanation of phenomena. This book explores mega-geomorphology. The parent science of geology has long emphasized large-scale features in its central discipline of tectonics. Although early proponents of largescale crustal mobilism, such as Alfred Wegener, were decidedly renounced by the mainstream scientific community, their ideas provided the stimulus for work that eventually transformed the Earth sciences.The plate tectonic model that emerged in the late 1960s was but a quantitatively geophysical confirmation of the elegant hypothesis developed by careful attention to large-scale structural patterns on the Earth´s surface.

Table: A Hierarchical Classification of Terrestria Geomorphological Features by Scale

Sl. No			**Approximate Time Scales of Persistence (years)**
1	107	Continents, ocean basins	10^8-10^9
2	10^6	Physiographic provinces, shieds, depositional plains	10^8
3	10^4	Medium-scale tectonic units (sedimentary basins, mountain massifs, domal uplifts)	10^7-10^8
4	10^2	Smaller tectonic units (fault blocks, volcanoes, troughs, sedimentary subbasins, individual mountain zones)	10^7
5	10-10^2	Large-scale erosional/depositional units (deltas, major valleys, piedmonts)	10^6
6	10^{-1}-10	Medium-scale erosional/depositional units or landforms (floodplains, alluvial fans, moraines, smaller valleys and canyons)	10^5-10^6
7	10^{-2}	Small-scale erosional/depositional units or landforms (ridges, terraces, sand dunes)	10^4-10^5
8	10^{-4}	Larger geomorphic process units (hillslopes, sections of stream channels)	10^3
9	10^{-6}	Medium-scale geomorphic process units (pools and riffles, river bars, solution pits)	10^2
10	10^{-8}	Microscale geomorphic process units (fluvial and eolian ripples glacial striations)	-

Of course, this is not intended to imply that microscale studies are unimportant in structural geology. Such studies tell much about the details of rock deformation and the fabric of resulting materials. The session here is that significant science occurs at all scales of study. Scientists neglect the study of one spatial scale to the peril of their advancement to understanding.

Scales of Study

Callieux and Romani believe that there are two major trends in modern geomorphology: towards quantification and towards more varied extensions. In the latter, they see an extension of concern to other scientific disciplines, to applied problems, to longer time scales, and to more ancient features. They also see an extension to greater spatial scales. The operative temporal and spatial scales of geologic phenomena span an immense range. Note that fluvial phenomena occupy an intermediate position in this scaling.

The various phenomena all follow the general equation:

$$S = aT^b$$

Where S is the size of the feature, T is the time and a and b are constants. The constant b is generally a scaling factor showing that big phenomena tend to last longer. The constant a seems to relate to the intensity of the process. Despite the utility of such scaling relationships, summarized in the concept of allometry problems arise at very large spatial and temporal extrapolations. However, perhaps more important than the details of scaling is the intellectual excitement of multiscale thinking. In his insightful review of geomorphological processes on planetary surfaces, Sharp observed that one of the sessions from the comparative study of landforms on different planets is to "think big."

The same can be said of the application of space-age global remote sensing technology to the study of terrestrial landforms. At the turn of the century, geomorphologists are once again ready to think big.

Time scales are also important in defining the nature of geomorphic investigations. The role of time scales in geomorphology is discussed more fully by Schumm and Lichty and Cullingford. The macroscale of geomorphic time is that over which major phases of erosion or deposition occur. These phases, which may be cyclic, are controlled by the geological processes of regional warping, mountain building, and crustal plate movement.

The scientific questions arising at the temporal macroscale concern the evolution of the planetary surface over millions of years. Many of the approaches to this time scale are historical, using elements of the landscape to reconstruct past events. Time scales are also important in defining the nature of geomorphic investigations. The role of time scales in geomorphology is discussed more fully by Schumm and Lichty and Cullingford et al.. The macroscale of geomorphic time is that over which major phases of erosion or deposition occur. These phases, which may be cyclic, are controlled by the geological processes of regional warping, mountain building, and crustal plate movement.

The mesoscale of geomorphic time is that which treats major changes in landforms and landscapes over hundreds to thousands of years. Examples include the growth and recession of glaciers, the aggravation and degradation of rivers, and the progradation and recession of shorelines. These changes generally involve a complex interplay between tectonic and climatic controls on geomorphological processes.

This is perhaps the most fertile area of geomorphic analysis, since it involves continual intellectual feedback between historical reconstruction and the study of modern processes.

The microscale of geomorphic time is that over which the major variables of tectonism and climate are assumed to be constant. The processes that characterize sand dunes, glaciers, rivers, or beaches are assumed to reflect only the short-term events that dictate local flow physics.

This is the temporal view of both the engineer and the process geomorphologist. Because of the ability to measure events that occur at the approximate scale of a year, this temporal scale is the richest source of quantitative geomorphic data. The ability to fashion predictions from studies at this scale also makes such studies immensely useful in application. People inhabit the dynamic surface of the Earth and must interact with its surficial processes. Reasoning from one scale to another is an instinctive mental exercise for scientists.

The geomorphologists of the 19th century generally perceived the details of landforms from foot, horseback, or even hands and knees. From such detail, they generalized to the larger synthesis of landscape form and processes. Nevertheless, they also realised the importance of a broad view. The great vistas of the Alps, Appalachians, and Colorado Plateau inspired regional assessment of landscape types.

The reasoning from distant views to the details of site specific investigation is perhaps no better illustrated than in G.K. Gilbert's study of the Henry Mountains in Utah. Gilbert's brilliant concept of the laccolith structure and its modification by erosional processes was conceived before he visited the scene! His field notes show that he conceived the relationships from distant views as he approached the mountains along the Waterpocket Fold.

UNIFORMITARIANISM AND NEOCATASTROPHISM

Uniformitarianism

Uniformitarianism, in geology, the doctrine suggesting that Earth's geologic processes acted in the same manner and with essentially the same intensity in the past as they do in the present and that such uniformity is sufficient to account for all geologic change. This principle is fundamental to geologic thinking and underlies the whole development of the science of geology.

When William Whewell, a University of Cambridge scholar, introduced the term in 1832, the prevailing view (called catastrophism) was that Earth had originated through supernatural means and had been affected by a series of catastrophic events such as the biblical Flood. In contrast to catastrophism, uniformitarianism postulates that phenomena displayed in rocks may be entirely accounted for by geologic processes that continue to operate—in other words, the present is the key to the past.

The expression *uniformitarianism*, however, has passed into history, because the argument between catastrophists and uniformitarians has largely died. Geology as an applied science draws on the other sciences, but geologic discovery had in the early 19th century outrun the physics and chemistry of the day. As geologic phenomena became understandable in terms of advancing physics, chemistry, and biology, the reality of the principle of uniformity as a major philosophical tenet of geology became established, and the controversy between catastrophists and uniformitarians largely ended.

Hutton's contributions

The idea that the laws that govern geologic processes have not changed during Earth's history was first expressed by Scottish geologist James Hutton, who in 1785 presented his ideas—later published in two volumes as *Theory of the Earth* (1795)—at meetings of the Royal Society of Edinburgh. Hutton showed that Earth had a long history that could be interpreted in terms of processes observed in the present. He showed, for instance, how soils were formed by the weathering of rocks and how layers of sediment accumulated on Earth's surface.

He also stated that there was no need of any preternatural cause to explain the geologic record. Hutton's proposal challenged the concept of a biblical Earth (with a history of some 6,000 years) that was created especially to be a home for human beings; the effect of his ideas on the learned world can be compared only the earlier revolution in thought brought about by Polish astronomer Nicolaus Copernicus, German astronomer Johannes Kepler, and Italian astronomer Galileo when they displaced the concept of a universe centred on Earth with the concept of a solar system centred on the Sun. Both advances challenged existing thought and were fiercely resisted for many years.

In *Principles of Geology*, 3 vol. (1830–33), Scottish geologist Sir Charles Lyell deciphered Earth's history by employing Huttonian principles and made available a host of new geologic evidence supporting the view that physical laws are permanent and that any form of supernaturalism could be rejected. Lyell's work in turn profoundly influenced English naturalist Charles Darwin, who recognized Lyell as having produced a revolution in science.

Lord Kelvin's contributions

The publication in 1859 of the conclusions of Darwin and British naturalist Alfred Russel Wallace on the origin of species extended the principle of uniformity to the plant and animal kingdoms. Although catastrophists continued to fight a rearguard action against the Huttonian-Lyellian-Darwinian view until the end of the 19th century, a new criticism was raised by William Thomson (later Lord Kelvin), one of the leading researchers on thermodynamics. Thomson pointed out that Earth loses heat by thermal conduction and that geologic processes may have changed as a consequence; he also concluded that this cooling placed an upper limit on Earth's age. With the discovery of radioactivity and the recognition that radioactive isotopes within the planet provide a continuing internal source of heat, it became clear that Thomson's conclusion that Earth was less than 100 million years old was incorrect, but his argument that Earth suffers an irreversible loss of energy remains valid.

Neocatastrophism

Neocatastrophism is the explanation of sudden extinctions in the palaeontological record by high magnitude, low frequency events (such as asteroid impacts, super-volcanic eruptions, supernova gamma ray bursts, etc.), as opposed to the more prevalent geomorphological thought which emphasises low magnitude, high frequency events.

OPENSYSTEM. ERGODIC PRINCIPLE

A promising new methodologic and conceptual framework is afforded by "general systems theory," which in its present state is less a theory than a collection of complementary concepts useful in treating complex macroscopic phenomena as an organic whole.

Ergodic Principle

In geomorphology, there are several discussions from theoretical approach and among these subjects; it can be emphasized on morphogenesis succession and its modeling in that subject. In different categories the course of ergodicity is one of the most important subjects. This filed which has been adopted of physical thermodynamic science also in geomorphology is under study. Simply the importance of this subject is referred to historical modeling, because one of the most imperative phenomena in biologic, geologic, and also geomorphologic fields is time and its relationship with changes and successions.

EQUILIBRIUM – TYPES OF EQUILIBRIA

Equilibrium is commonly used to describe geomorphic systems that can adjust to changes by reaching a steady state. As a conceptual framework, equilibrium emphasizes the relation between present form and process. Systems analysis is a conceptual framework that can be used to quantify and relate material or energy flows that comprise a geomorphic system. Despite the apparent similarity between the language of systems analysis and the descriptive examples of geomorphic equilibrium, these two conceptual frameworks are divergent in many respects. From the vantage point of a systems approach in geomorphology, many types of equilibrium purported to describe geomorphic systems, such as dynamic equilibrium and graded equilibrium, are seen to be very different from one another and in turn very different from the kinds of equilibrium resulting from systems analysis.

Equilibrium is a single word that embraces multiple concepts. The largely qualitative use of equilibrium within geomorphology has fostered imprecision and even outright error; as a result the term, for many, has degenerated to the status of a non-corrigible regulative principle. Although a few geomorphologists make precise use of equilibria terminology, their precision is invariably lost on the larger audience.

Equilibrium is associated with force in dynamics, with energy in thermodynamics (and probability by extension), and with pure numerical behavior in mathematics. In General System Theory, equilibrium is derived from thermodynamics but applied, by analogy, almost exclusively to mass. In addition, a purely geomorphic version of equilibrium (dynamic equilibrium) stems from the work of G. K. Gilbert (1876; 1877) and is based on sediment flux at the basin scale. Unfortunately, Gilbert's concept/term has been distorted by some geomorphologists in their attempt to establish linkage between equilibria concepts, especially between those relating to energy and mass.

Entropy and equilibrium

- the concept of entropy, the degree of disorder in an isolated system, and the second law of thermodynamics dictate that, irrespective of the initial level of energy, energy becomes more evenly distributed over time (*i.e.*, energy gradients are reduced) and processes proceed at an ever decreasing rate
- ultimately, energy is evenly distributed, entropy is a maximum and free energy (available for conversion to work) is zero, this is, the state of equilibrium where the most probable energy distribution is uniform with least work (least landscape change)
- these concepts explain characteristic (least-work, graded) landforms like longitudinal stream profiles, stream channel meanders, dendritic drainage networks, and concave-conves slope profiles
- entropy is maximized when all possible energy states have equal probability (*i.e.*, random distribution)
- thus there is rapid change away from any highly abnormal configuration of a system, corresponding to the notion of initially rapid change in response to an input or disturbance and then a decreasing rate of change over time

However geomorphic systems are also subject to order imposed by structural (geologic) control and to repeated and variable inputs, suggesting various kinds of equilibrium:

1. static
   - o some system properties are unchanging, absolutely and relatively, over a period of time
2. stable
   - o tendency to revert to a previous condition after a limited disturbance
3. unstable
   - o a small disturbance results in continuous movement away from an old equilibrium and towards a new stable one
4. metastable
   - o an incremental change (trigger mechanism) pushes a system across a threshold from stable equilibrium to a new equilibrium
5. steady state
   - o numerous small-scale about a mean that has no trend
6. thermodynamic
   - o a tendency towards maximum equilibrium, second law of thermodynamics
7. dynamic
   - o balanced fluctuations about a trending, nonrepetative mean value; often called quasi-equilibrium because of the tendency towards a steady state with a trending mean (*i.e.*, equilibrium is never attained)
8. dynamic metastable
   - o both dynamic and metastable, fluctuations about a trending mean interspersed by large jumps as thresholds are crossed; used by many geomorphologists to explain landscape behaviour

5

Complex Response and Geomorphic Thresholds

Geomorphology has inherited the idea of systems from physics, biology, and, to a lesser extent, chemistry. Physicists recognize three main kinds of system: simple systems, complex but disorganized systems, and complex and organized systems. The first two conceptions of systems have a long and illustrious history of study. Since at least the sixteenth-century revolution in science, astronomers have referred to a set of heavenly bodies connected together and acting on each other in accordance with certain laws as a 'simple' system: the solar system is the Sun and its planets; the Jovian system is the planet Jupiter and its moons. In geomorphology, a few boulders resting on a hillslope is a simple system. The conditions required to dislodge the boulders and their fate once they have been dislodged are predictable from mechanical laws involving forces, resistances, and equations of motion, in much the same way that the motion of planets around the Sun can be predicted from Newtonian laws.

In the complex but disorganized conception of systems, a vast number of objects interact in a weak and haphazard manner. An example is gas in a jar. This system could consist of more than 10^{23} molecules colliding with each other. In the same way, the countless individual particles in a hillslope mantle are viewable as a complex but somewhat disorganized system, even if the hillslope mantle as a whole does have an organization. In both the gas and the hillslope mantle, the interactions are rather haphazard and far too numerous to study individually, so aggregate measures must be employed.

In a more modern notion of systems, which has its roots in biology and ecology, objects interact strongly with one another to form systems with a complex and organized nature. Most biological systems and ecosystems are of this kind, but many other structures at the Earth's surface display high degrees of regularity and rich connexions and may be thought of as complexly organized systems. Hillslopes, rivers, and beaches are examples. This kind of system, and its extension to non-equilibrium cases, has dominated systems thinking in geomorphology since the 1970s, although earlier systems ideas are still important.

COMPLEX RESPONSE

A complex response refers to an environmental reaction to change that occurs at multiple levels to multiple objects, and can induce a chain reaction of responses to a single initial change. It is akin to the butterfly effect: one small event (change) can cascade through a given system creating new agents of change, and operating at several levels. The term is most commonly used in fluvial geomorphology, or the study of river systems and changes within those systems.

Systems of complex disorder and classical thermodynamics

8Nineteenth-century physicists invented classical thermodynamics, which is the study of heat transformation and exchange. Classical thermodynamics involves the study of heat (and so with the collision and interaction of particles) in large and closed systems in, or near, equilibrium states. A large system in classical thermodynamics contains a huge number of interacting particles (for example, colliding molecules in a gas or a huge number of billiard balls on a big billiard table for instance).

Newtonian dynamics cannot tackle systems of this complexity because there are too many equations to handle. In the late nineteenth century, Boltzmann and Gibbs showed that systems that are complex but disorganized can be studied using a new kind of mathematics to find certain quantities of interest. The new mathematics was statistical averaging or, more technically, entropy maximizing (Wilson, 1981, p. 39). These methods cannot predict the behaviour of any one particle of gas, or one billiard ball among many. However, they can predict the distribution of particles among energy states and aggregate measures such as the pressure and temperature of a gas, and the distribution of velocities of billiard balls and the average rate at which a ball strikes a cushion. The predictions made by this branch of classical thermodynamics apply to closed system at, or very near to, equilibrium.

Energy and mass conservation in systems

Regarding geomorphic systems as systems of complex disorder opens up an interesting and powerful line of enquiry that applies principles of energy and materials accounting. Although the principles of accounting are simple, they can lead to sophisticated methods such as input–output analysis. They can also lead to statements of energy conservation and mass conservation that, in conjunction with principles of classical dynamics, assist in the study a variety of geomorphic systems. To be sure, the laws of Newtonian dynamics and thermodynamics are widely used in geomorphology. These laws cover four kinds of basic equations: balance equations, physical–chemical state equations, phenomenological equations, and entropy balance equations (Isermann, 1975).

Systems of complex order

The thermodynamics of open systems

A closed system in classical thermodynamics does not exchange energy or material with it surroundings. Some authorities define a closed system as one that exchanges energy but not matter with its environment and an isolated system as one that exchanges neither energy nor matter with it environment. In closed systems, the total amount of energy, *E*, is always conserved (as stipulated by the first law of thermodynamics), the amount of available energy inevitably decreases to zero (as dictated by the second law of thermodynamics), and the entropy, *S*, of the system (the amount of unusable energy) increases to a maximum. Around the middle of the twentieth century, the theory of irreversible processes and open systems, which physicists, chemists, and biologists developed, led to a new thermodynamics. Open systems exchange energy or matter (or both) with their surroundings and can exhibit nonlinear behaviour. The immensely potent idea of open systems was the brainchild of von Bertalanffy (1932). From about 1932, von Bertalanffy explored the implications of viewing organisms as open systems, and, building on the groundbreaking work of Lotka (1924, 1954), which drew on chemical reaction theory, couched the dynamics of biological systems in terms of simultaneous differential equations (e.g. Bertalanffy, 1950). This work was the inspiration for the eventual injection of open systems concepts into geomorphology.

GEOMORPHIC THRESHOLDS

Rivers are subject to thresholds of several types that define significant changes in processes and morphology and delimit distinctive riverine landscapes and habitats. Thresholds are set by the conditions that govern river channel process and form, amongst which the most important are the flow regime, the quantity and calibre of sediment delivered to the channel, and the topographic setting (which determines the gradient of the channel). These factors determine the sediment transport regime and the character of alluvial deposits along the channel.

Thresholds

Thresholds have been recognized in many fields and their importance in geography has been discussed in detail by Brunet (1968). Perhaps the best known to geologists are the threshold velocities that are required to set in motion sediment particles of a given size. With a continuous increase in velocity, threshold velocities are encountered at which movement begins, and with a progressive decrease in velocity, threshold velocities are encountered at which movement ceases. These are Brunet's (1968, pp. 14 and 15) 'thresholds of manifestation' and 'thresholds of extinction', and they are the most common types of thresholds encountered. However, when a third variable is involved, Brunet (1969, p. 19) identified 'thresholds of reversal.' An example of this type of threshold is Hjulstrom's (I935) curve showing the velocity required for movement of sediment of a given size. The curve shows that velocity decreases with particle size until cohesive forces become significant, and then the critical velocity increases with decreasing grain size. Another example of this type of relationship is the Langbein-Schumm (1958) curve which shows sediment yield as directly related to annual precipitation and run-off until vegetation cover increases sufficiently to retard erosion. At this point there is a decrease in sediment yield with increased run-off and precipitation. Perhaps thresholds is not a good word to describe the critical zones within which these changes occur, but it is a simple and easily understood term.

The best known thresholds in hydraulics are described by the Froude and the Reynolds numbers, which define the conditions at which flow becomes supercritical and turbulent. Particularly dramatic are the changes in bed-form characteristics at threshold values of stream power. In the examples cited, an external variable changes progressively, thereby triggering abrupt changes or failure within the affected system. Response of a system to an external influence occurs at what will be referred to as extrinsic thresholds. That is, the threshold exists within the system, but it will not be crossed and change will not occur without the influence of an external variable.

Thresholds of the other type are intrinsic thresholds, and changes occur without a change in an external variable. An example is long-term progressive weathering that reduces the strength of slope materials until eventually there is slope adjustment (Carson, 1971) and mass movement (Kirkby, 1973). Following failure, a long period of preparation ensues before failure can occur again (Tricart, 1965, p. 99). Glacial surges, that are not the result of climatic fluctuations or tectonics (Meier and Post, 1969) probably reflect periodic storage and release of ice, as an intrinsic threshold of glacial stability is exceeded. In semi-arid regions sediment storage progressively increases the slope of the valley floor until failure occurs by gullying.

This is a special type of intrinsic threshold, the geomorphic threshold (Schumm, 1973). It is a result of landform (slope) change through time to a condition of incipient instability and then failure. Another example of a geomorphic threshold is provided by Koons (I955) who described morphologic changes resulting in the collapse of sandstone-capped shale cliffs in the Mesa Verde region in south-western Colorad6. Beneath a vertical cliff of Mesa Verde sandstone is a 32 to 38 degree slope of weak Mancos Shale. Through time, the basal shale slope is eroded and reduced in height, thereby producing a vertical shale cliff beneath the sandstone cap. At some critical height the cliff collapses and the cycle begins again. The episodic retreat of this escarpment is the result of the change in cliff morphology under constant climatic, base level and tectonic conditions

Intrinsic thresholds are probably common in geologic systems, but only geomorphic examples will be considered here. A geomorphic threshold is one that is inherent in the manner of landform change; it is a threshold that is developed within the geomorphic system by changes in the morphology of the landform itself through time. It is the change in the landform itself that is most important, because until it has evolved to a critical situation, adjustment or failure will not occur. The concept of geomorphic thresholds, which involves landform change without a change in external controls challenges the well-established basis of geomorphology, that landform change is the result of some climatic, base-level or land-use change. Therefore, the significance of the geomorphic threshold concept for the geomorphologist is that it makes him aware that abrupt erosional and depositional changes can be inherent in the normal development of a landscape and that a change in an external variable is not always required for a geomorphic threshold to be exceeded and for a significant geomorphic event to ensue.

It is important to stress again that the geomorphic threshold as defined above is an intrinsic threshold. If as a result of a climate change, discharge and flow velocities in a stream channel are increased, the resulting bank erosion and meander cut-offs may convert a meandering stream to a braided channel.

In this case the cause of the pattern conversion is extrinsic, but if, as a result of increasing channel sinuosity under unchanging average discharge, cut-offs occur which convert a meandering reach to a straight reach the control is intrinsic. In the first case the change may be regional in nature and in the second it may be local.

However, explanation of locally anomalous, erosional or depositional features is what is required for rational landform management. During recent discussions with students and colleagues, it became apparent that the term geomorphic threshold is being used in a much broader sense than originally intended, and a revision of the definition is required.

The concept of intrinsic geomorphic controls was stressed because it initially provided a new approach to the understanding of the details of the landscape that could be used for predictive purposes. Nevertheless, there can also be extrinsic geomorphic thresholds.

For example, in common usage 'thresholds' can be the result of either cause or effect. That is, we speak of hydraulic, velocity, shear, and stream power thresholds above which sediment moves or banks fail, but we can also speak of bank, channel and slope stability thresholds, when the forces causing the failures are not clearly identified and understood. Therefore, geomorphic thresholds can be of two types, and they can be redefined in the following way. A geomorphic threshold is a threshold of landform stability that is exceeded either by intrinsic change of the landform itself, or by a progressive change of an external variable. Although the original definition is broadened, the concept of abrupt landform change remains

Evidence for geomorphic thresholds

Recent field and experimental work supports the concept of geomorphic thresholds. The best examples result from investigations of gully distribution and stream pattern variability

Within this area, valleys were selected in which discontinuous gullies were present. The drainage area above each gully was measured on maps, and critical valley slopes at the point of gully development were surveyed in the field. No hydrologic records exist, so drainage-basin area was selected as a substitute for run-off and flood discharge.

When this critical valley slope is plotted against drainage area, the relationship is inverse, with gentler slopes being characteristic of large drainage areas. As a basis for comparison, similar measurements were made in ungullied valleys, and these data. The lower range of critical slopes of the unstable valleys coincides with the higher range of slopes of the stable valleys. In other words, for a given drainage area it is possible to define a critical valley slope above which the valley floor is unstable

Tide-dominated systems

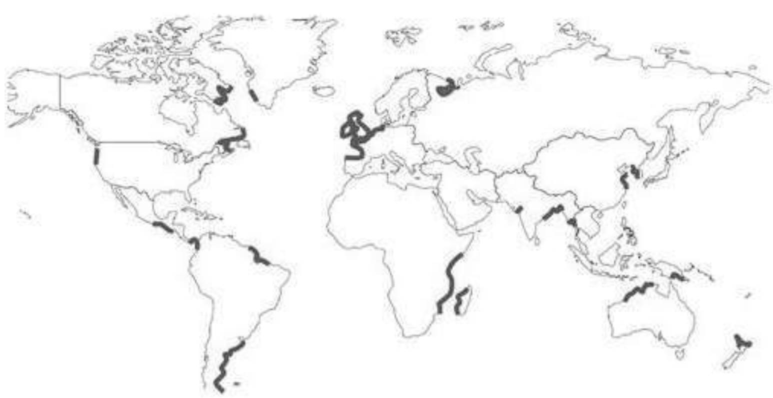

Figure : World distribution of macrotidal coastal zones, where the maximum tidal range exceeds 4 m.

Tide-dominated estuarine systems occur on coasts with strong semi-diurnal tides. These estuaries are characterised by a wide mouth; estuarine channels are narrowing inland. Examples are the Charente and the Hooghly estuaries. A mouth barrier is absent or reworked into a submarine ebb-tidal delta.

Channels are scoured into old river sediments or recent marine deposits, a process that has mainly taken place over the past millennia when the marine transgression has stabilized close to current sea level. At the head of the estuary a single main channel connects to the river upstream.

Tidal flats are present along the channels; they are mostly built with sediment imported from the sea. An example is Gomso Bay. The highest landward parts of tidal flats develop into vegetated salt marshes, unless this is prevented by frequent strong wave action. These marshes are flooded only at high water. The coarsest sediments are found at the estuarine mouth, where strong tide- and wave-induced currents act on the bed. Fine sediments are deposited on the tidal flats and in the inland channels. Most of the fine sediments in the upper estuarine zone are of fluvial origin.

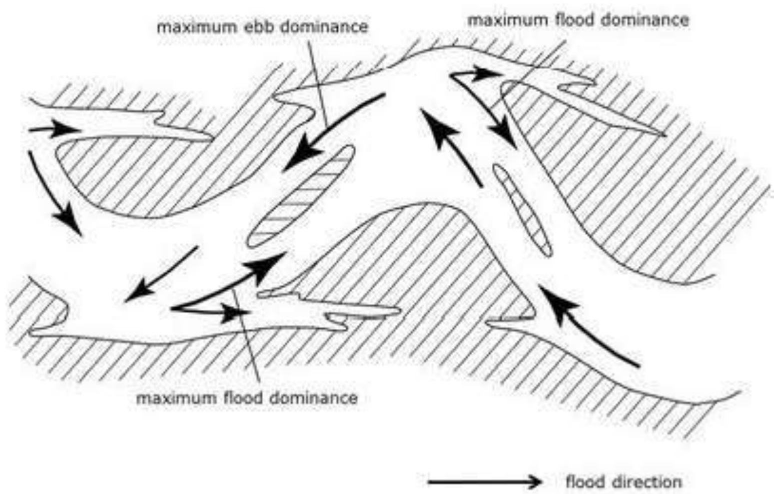

Figure : Schematic picture of a flood and ebb dominant flow channels in a meandering channel system.

Channels are deeply scoured if the channel bed is sufficiently erodible, but they can remain shallow if the substrate is very hard or if the substrate, on the other hand, consists of very mobile non-cohesive fine sediment.

Tidal channels are typically meandering, which is the result of non-linear interaction between morphology and flow dynamics - due in particular to the "overshoot" of the flow at channel bends (so-called inertial delay). The same phenomenon also makes that flood and ebb flows tend to follow different paths through the estuary, eventually leading to a multi-channel system of flood and ebb dominated channels .

In places where flood and ebb channels intersect, sand transport converges. This creates thresholds, which form an obstacle to shipping. In order to prevent nuisance to shipping, these thresholds must be dredged regularly. An example is the Western Scheldt. Lateral migration of meandering channels occurs in estuaries with strong tides if the channel bed consists of sufficiently mobile sediments. Fast channel migration is observed, for example, in the mega-tidal Baie du Mont Saint Michel . Migrating channels can reshape the estuarine landscape in a quasi-cyclic process by their strong lateral scouring power.

The tidal wave is distorted when propagating up-estuary. Higher harmonic components are generated with frequencies that are multiples of the frequency of the dominant semi-diurnal tidal frequency M2.

The harmonic M4 tide is particularly relevant, because it generates flood currents of different strength compared to ebb currents. For strongly converging (upstream narrowing) estuaries the tidal amplitude increases; after reaching a maximum the tidal amplitude decreases further upstream where the fluvial influence becomes substantial.

In most tide-dominated estuaries the high-water (HW) crest of the tidal wave travels faster than the low-water (LW) trough, due to the greater water depth at HW compared to LW. In this case an asymmetry develops with shorter flood duration and longer ebb duration. Therefore, flood currents are generally stronger than ebb currents in the lower part of the estuary; this favours import of marine sediment. Estuarine circulation, driven by salinity gradients, also contributes to sediment import. Both processes contribute to the development of a high turbidity zone in the estuary, the so-called estuarine turbidity maximum. However, under conditions of high river discharge, sediment export dominates over import. Tide-induced import and river-induced export both contribute to the net long-term sediment balance of the estuary . A more detailed discussion of the various estuarine transport processes is given in the articles Tidal asymmetry and tidal basin morphodynamics and Estuarine circulation.

6

Space and Time in Geomorphology

TIME AND SPACE IN GEOMORPHOLOGY

- time and space are the two fundamental perspectives of all human enquiry and experience
- the disciplines of history and geography are characterized not by subject matter but by the study of temporal and spatial relations, respectively
- geomorphology has strong ties to geology historically and topically, so there is a strong theme of earth history, especially since we focus on retrodiction rather than prediction and landforms evolve at geological time scales
- however given the affiliation with geography and relation to climate, hydrology, biogeography and soils, there is also a significant spatial element to the study of landforms

Fundamental Concepts of Landscape Evolution

Uniformitarianism:

- introduced by James Hutton (1788), champoined by John Playfair and named by Charles Lyell (1830) in "Principles of Geology"

- usually defined by the phrase "the present is the key to the past" but also encompasses the notions that
  - physical and chemical laws are invariant with time,
  - geomorphic processes are of the same kind and rate as in the past, and thus
  - the history of the earth can be explained from current observations and relationships (causes)
- uniformitarianism was such a profound departure from the prevailing view of nature that it developed a broad scope in terms of its implications and application
- originally Hutton was seeking an alternative to supernatural cataclysmic origin of a young earth, but through the work of proselytes like Playfair and Lyell it evolved into a complex idea with two components.

Substantive uniformitarianism:

- the underlying hypothesis which postulates the uniformity of material conditions and rates of processes; the notion of temporal uniformity of material conditions and process rates (gradualism) can be dismissed;
- "The picture that physics gives us of the universe is of relative macrocosmic stability overlying violent and discontinuous microcosmic activity... Uniformitarianism of this sort defies the findings and whole tenor of modern science."

Methodological uniformitarianism:

- all historical enquiry assumes that natural laws are constant in time and space and that "no hypothetical and unknown process can be invoked if observed historical results can be explained by presently observable processes"
- however, earth physics may change over earth history with changes in the composition of the atmosphere or position of the earth in the universe; but natural laws

have a restricted domain beyond which they don't apply nor can they be applied since these are "highly inaccessible" domains

- a more significant objection to methodological uniformitarianism as a formal concept is that it lacks any contemporary usefulness because the assumption of constancy of nature is so fundamental to all modern scientific enquiry that it does not require a formal statement; there is no alternative mode of thinking as in Hutton's day nor does anyone even consider an alternative
- science demands generalizations, the simpler the better; even if there is no pattern to the frequency and magnitude of natural events, at the very least we have to seek approximations of constancy in "the rate of change of rate of change" so that are laws describe the behaviour of nature; thus "the principle of uniformity dissolves into a principle of simplicity that is not peculiar to geology but pervades all science and even everyday life".

Neocatastrophism:

- the contemporary alternative to substantive uniformitarianism:
- neocatastrophism is the modern version of the notion that rare, high-magnitude events have played an important role in landscape evolution
- it is devoid of the original constraints of a short geological record and appeal to supernatural powers, *i.e.*, it is scientifically-based whereas the original catastrophism that uniformitarianism overcame was biblically based
- ironically, to suggest neocatastrophism in the hay day of uniformitarianism was almost as heretical as the views of Steno and Hutton
- Lyell efforts to popularize uniformitarianism, plus the concept of evolution advocated by his close friend Charles

Darwin, served to entrench the concept of slow change in the earth and natural sciences, defeating the notions of a short earth and supernatural powers

- in geomorphology, catastrophism was replaced by gradualism, the idea that landscape change is slow, occurs in small increments, and involves weak forces; the dismissal of rare, high-magnitude events was as much emotional as scientific and accounts for the initial response to Bretz's (1923) explanation of the Channeled Scablands of eastern Washington State during the 1920s and 30s
- acceptance of the importance of large events occurred as paleontologists supplied evidence of mass extinctions and other geologists described rocks that show evidence of large, individual geomorphic events in sedimentary rocks
- for most geomorphologists, neocatastrophism represents only reappraisal of the importance of geomorphic processes in the terms of magnitude and frequency if there is room for the interpretation of unique events, however, then geology and geomorphology become a form of history rather than science.

Erodicity (ergodic theorem, hypothesis, assumption):

- because most landscapechange occurs over longer than human timescales, geomorphologists infer the nature of landscape evolution by comparing similar landforms of different age, that is at different stages of evolution
- that is, a temporal sequences (landform evolution) cannot generally be observed and thus must be reconstructed from a theoretical basis (*i.e.* by invoking ergodicity)
- this practice is so common that we are not always conscious of the underlying assumption, that is, that we can substitute observation in space for observations in time
- despite its significance, the ergodic assumption receives little attention in the geomorphic literature
- the concept of ergodicity developed in physics because molecules move so rapidly compared to the duration of

observation that their spatial distribution cannot be easily established; therefore physicists hypothesized that "the mean observations of an individual made over time is equal to the mean observations of many individuals made of may individuals at a single moment in time over an area"

- thus the ergodic assumption in physics was formulated so that observations at different times could be used as a surrogate for the spatial distribution at a single moment, but in geomorphology
  - this hypothesis has never been tested (*i.e.* studied quantitatively), rather ergodicity is assumed
  - space is used as a surrogate for time,
  - ergodicity is not applied in the original sense of time and space 'averages', and
  - but is used to translate spatial distribution into temporal sequences, which was not intended by the original concept
- since the use of ergodicity in geomorphology is conceptual rather than mathematical and departs from the formal concept, phrases like "location-for-time" or "space-for-time" substitution are used to describe the assumption in geomorphology
- although there are a few examples of truly ergodic modelling in fluvial geomorphology; use of space-for-time substitution has been applied most often to studies of scarp evolution; scarps tend to form quickly (*e.g.* tectonism, mass wasting) or undercutting of erosional scarps ceases as streams shift or relative sea level drops
- a special case of ergodicity is allometry, the study of proportional changes correlated with variations in size
  - dynamic allometry describes changes in relationships between parts of an individual through time; static allometry describes changes in relationships between variously-sized members of a group at a moment in

time; both kinds of relationships are usually expressed as power functions

- allometry has been used to describe the evolution of cirques and alluvial fans
- comparisons of the dynamic and static allometry of the same landform reveal that the relationships differ (analogous to differences in hydraulic geometry along a river at the same time versus at a station at different times); a statistically significant difference would violate the ergodic assumption

In summary, there must be some theoretical foundation for the widespread assumption that the spatial distribution of landforms reveals something about the evolution of individual landforms or, that is, that present landscapes can reveal something about their paleo-equivalents (methodological uniformitarianism). Ergodicity would seem to be this framework, however it has applied to geomorphology without rigour or formal definition.

TIME – CYCLIC, GRADED, STEADY

Distinctions between cause and effect in landform development depend on the span of time involved and the size of the geomorphic system under consideration. Depending upon the temporal and areal frame of reference, variables such as channel morphology may be either dependent or independent. In terms of geologic time, landforms represent a stage in an erosion cycle and are dependent on time. On a short-term basis, components of geomorphic systems may be regarded as systems in dynamic equilibrium or in a steady state and are independent of time.

Time

The early karst studies of Cvijic and Grund at the turn of the 20th century were strongly influenced by the cyclic evolutionary teachings of Davis and Penck. Thus, karst topography has been thought to evolve through stages of development beginning with fluvial action and the initial formation of dolines. As the surface drainage is slowly captured by swallow holes and the developing subsurface drainage, dolines coalesce to form uvalas, which, in turn, expand to poljes. This sequence is now considered to be antiquated, and just as to Roglic, poljes are structurally controlled and are not related to uvalas.

It is questionable whether any two karst regions anywhere in the world have evolved sequentially in the same manner or even display identical morphological forms. Each karst region evolves within its unique combination of dynamic and static factors. However, this does not preclude the comparison of areas of similar morphology in an attempt to evaluate common factors. From the examples of Landsat imagery in this stage, it becomes apparent that each region has developed its own morphologic signature. To understand the evolution of a given karst terrain, it is necessary to comprehend the roles of length of time, style, and intensity of the energy flow through the hydrologic system. Those karst regions with the largest potential energy, related to the thickness of the vadose zone, and kinetic energy, or flooding, will also exhibit a higher energy morphology more rapidly and more fully than regions of lower total energy. High-energy morphology can be manifested in the form of larger caves, increased relief, density, and magnitude of dolines, hums, flood deposits, and a more extensive underground solution network.

Karst from Space

Within the realm of geomorphology, each geologic process leaves its own imprint on the Earth´s landscape, and each process will develop its own characteristic assemblage of landforms. For karst landscapes, this imprint is expressed as solution morphology on the regional scale. Because of the relatively large ground resolution distance of 79 m, many karst landforms cannot be discerned from Landsat MSS images; only the larger solution and fracture-controlled dolines and uvalas can be recognized.

Thus, the advantage of the space perspective is not the identification or the recognition of individual landforms, but the collective pattern and texture they impart to a region of hundreds or thousands of square kilometers. This can occur only where given geologic processes and materials, such as carbonates, dominate a large region for a sufficient length of time. The Landsat imagery discussed in this stage has been chosen from the known karst regions of the world. Some well-researched regions are not included either because good imagery has not been available or because man´s overprint is so dominant as to obscure natural topographic patterns.

A cursory examination of all the karst images reveals a number of features specifically associated with the solution process:

- Lack of drainage patterns
- Dimpled pock-marked texture
- Fracture sets with developed relief
- Uniformly scattered residual karst hills
- Circular basins or lakes
- Large flat karst plains and poljes

Most important is the absence of a well-developed integrated surface drainage. Indeed, for the Nullarbor Plain, Australia, and Yucatan, Mexico, it is difficult to identify any recognizable drainage. Major rivers can be seen to incise the south China carbonate plateaus, but there is a distinct lack of well-developed tributaries draining into the river channels.

Most karst areas exhibit a relatively homogeneous pock-marked or dimpled texture. In south China, this pattern is overprinted with solution-incised joint sets. In Florida, the pattern is expressed as numerous large water-filled collapsed basins. Joint sets are not recognized in all karst areas at Landsat scale.

This does not preclude the presence of joints, but merely points out that they may be obscured because:

- Joints are too closely spaced,
- Solution relief of joints is slight,
- Dense vegetation cover is present, or
- Joints may indeed be absent or sparse.

What accounts for the large differences observed in karst regions is not variation in the karst process, but variation in lithology, structure, and the overprint of other processes like fluvial, tectonic, eolian, and glacial, and variations in the energy flow through the karst system through time.

Geologic Time Scale

The standard geologic column is the basis for the *geologic time scale*. The names applied to the systems of rocks are also used for the *periods of time* during which the respective systems of rocks were deposited. Thus, we use the term *Carboniferous Period* for the time during which the *Carboniferous System*—the Mountain Limestone and Coal Measures of Smith—were laid down. Appended to it are remarks as to the sources of the names and the ages of several minerals as found from studies of radioactivity. From these ages in years it is apparent that more subdivisions are recognized among younger rocks than among older. Just as in human history, the documents are more numerous and the gaps in knowledge less wide the more recent the period with which we deal.

Gaps in the Time Scale

Divisions of the standard column are based on abrupt changes in the fossil assemblages of the strata in Europe. These points of division naturally were those marking the longer intervals of erosion or nondeposition in the European stratigraphic series. The longer the interval lost by erosion, the more prominent the fossil differences would be. As stratigraphic work was extended to other continents, however, fossil assemblages intermediate between those of adjacent periods as recognized in western Europe were discovered.

With more and more stratigraphic work the gaps in the type sections have continued to decrease. Beds containing fossil assemblages that pose "boundary problems"—that is, uncertainty as to their correlation with the upper part of one European system or the lower part of the next higher—are thus present at one place or another for nearly every systemic boundary.

Even the boundary between the Paleozoic and Mesozoic Eras, long thought to represent an interval during which none of the present continents received deposits, is apparently bridged by an essentially complete succession of beds in the Himalayan Mountains. Even in Europe, where the divisions were first established, further research has served to narrow the lost gaps in the record. For example, a new division, the Paleocene, has been recently established for strata intermediate in age between Eocene and Cretaceous, where a long gap had been thought to exist.

Now we find that different paleontologists refer the same strata (let us say, in Wyoming) to the Cretaceous or to the Paleocene, depending on their respective appraisals of the relative affinities of the fossils of these beds to that of one or another of the European sections.

Intercontinental Correlation

Thus, although the original divisions of the type section were based on stratigraphic relations of the beds, our correlations of distant strata are necessarily based on a *comparison of fossils*, not strata. The demonstration of an interruption in stratigraphic succession in the type area can give no help in correlating beds in a distant part of the world. A stratigraphic break in Nevada, for example, cannot be used as evidence that the beds above it are Permian and those below Carboniferous. This assignment of age to the Nevada strata can be made only on the basis of comparison of fossils with those of the type section of Europe. The uncertainties posed by "boundary problems," however, are concerned largely with form instead of fact. Such disputed strata, it is generally agreed, must represent times between, or at least partially bridging, the gaps in the type areas of Europe. Further detailed stratigraphic work cannot but fill other parts of the gaps. This is indeed a tribute to the precision of correlation by fossils rather than a criticism of its uncertainties.

The Duration of Geologic Time

With the building of the geologic time scale there came a growing appreciation of the vast duration of geologic time. Contemplation of even the time required to lay down the 500-foot thickness of chalk in the Paris Basin is upsetting to the man who thinks of time only in terms of the span of human life. Some strata of the chalk are composed of the skeletons of minute animals and plants. Similar deposits are accumulating today at rates so low as to defy measurement—certainly no more than a few inches per century and probably much less. Yet the chalk represents only a small part of Cretaceous time, and the Cretaceous itself is a mere fraction of geologic time.

Not only the vast thickness of sediment impressed geologists with the immensity of geologic time, but also the vast parade of life recorded by the fossils—the development of thousands upon thousands of new species with advancing time and the gradual dying out of whole fossil assemblages and their succession by others—time after time—can only be conceived of as involving millions of years, unless greatly different rates of change prevailed in the past.

Time Represented by the Geologic Column

Spurred by their curiosity to translate geologic time into years, geologists tried many approaches. The simplest and most obvious one was to add the greatest thickness of strata of each period and thereby get the total thickness for all geologic time. This total, divided by the present annual rate of sedimentation, might be thought to give the length of geologic time in years.

However, so many assumptions are involved that such a calculation becomes meaningless. How can we arrive at an average annual rate of sedimentation when we know that it may take a century to lay down less than an inch of chalk, whereas a desert cloudburst may deposit 40 feet of gravel, sand, and boulders overnight?

And how can we measure the time of scour and nondeposition in marine sediments when we know that a single storm may remove the accumulation of years from part of the sea bottom and perhaps pile up more sand on a nearby beach than was deposited during the whole preceding decade? In short, the rate of sedimentation varies so greatly and has been so little measured that the average annual rate can only be guessed at. This method can only yield an order of magnitude for geologic time. By such a calculation, the British geologist Sollas estimated (1899) that 34 to 75 million years have passed since the beginning of the Paleozoic —the larger figure including a guess at the length of the lost intervals in the record. These figures have a certain interest but surely are not reliable.

MAGNITUDE AND FREQUENCY

Intensity of earth quaking and magnitude of earthquakes

Quaking or shaking of the earth is a common phenomenon undoubtedly known to humans from earliest times. Prior to the development of strong-motion accelerometers that can measure peak ground speed and acceleration directly, the intensity of the earth-shaking was estimated on the basis of the observed effects, as categorized on various seismic intensity scales. Only in the last century has the source of such shaking been identified as ruptures in the Earth's crust, with the intensity of shaking at any locality dependent not only on the local ground conditions but also on the strength or *magnitude* of the rupture, and on its distance.

The first scale for measuring earthquake magnitudes was developed by Charles F. Richter in 1935. Subsequent scales have retained a key feature, where each unit represents a ten-fold difference in the amplitude of the ground shaking and a 32-fold difference in energy. Subsequent scales are also adjusted to have approximately the same numeric value within the limits of the scale.

Although the mass media commonly reports earthquake magnitudes as "Richter magnitude" or "Richter scale", standard practice by most seismological authorities is to express an earthquake's strength on the moment magnitude scale, which is based on the actual energy released by an earthquake.

Frequency of occurrence

It is estimated that around 500,000 earthquakes occur each year, detectable with current instrumentation. About 100,000 of these can be felt. Minor earthquakes occur nearly constantly around the world in places like California and Alaska in the U.S., as well as in El Salvador, Mexico, Guatemala, Chile, Peru, Indonesia, Philippines, Iran, Pakistan, the Azores in Portugal, Turkey, New Zealand, Greece, Italy, India, Nepal and Japan. Larger earthquakes occur less frequently, the relationship being exponential; for example, roughly ten times as many earthquakes larger than magnitude 4 occur in a particular time period than earthquakes larger than magnitude 5. In the (low seismicity) United Kingdom, for example, it has been calculated that the average recurrences are: an earthquake of 3.7–4.6 every year, an earthquake of 4.7–5.5 every 10 years, and an earthquake of 5.6 or larger every 100 years. This is an example of the Gutenberg–Richter law.

The number of seismic stations has increased from about 350 in 1931 to many thousands today. As a result, many more earthquakes are reported than in the past, but this is because of the vast improvement in instrumentation, rather than an increase in the number of earthquakes. The United States Geological Survey estimates that, since 1900, there have been an average of 18 major earthquakes (magnitude 7.0–7.9) and one great earthquake (magnitude 8.0 or greater) per year, and that this average has been relatively stable. In recent years, the number of major earthquakes per year has decreased, though this is probably a statistical fluctuation rather than a systematic trend. More detailed statistics on the size and frequency of earthquakes is available from the United States Geological Survey (USGS). A recent increase in the number of major earthquakes has been noted, which could be explained by a cyclical pattern of periods of intense tectonic activity, interspersed with longer periods of low intensity. However, accurate recordings of earthquakes only began in the early 1900s, so it is too early to categorically state that this is the case.

Most of the world's earthquakes (90%, and 81% of the largest) take place in the 40,000-kilometre-long (25,000 mi), horseshoe-shaped zone called the circum-Pacific seismic belt, known as the Pacific Ring of Fire, which for the most part bounds the Pacific Plate. Massive earthquakes tend to occur along other plate boundaries too, such as along the Himalayan Mountains. With the rapid growth of mega-cities such as Mexico City, Tokyo and Tehran in areas of high seismic risk, some seismologists are warning that a single quake may claim the lives of up to three million people.

Induced seismicity

While most earthquakes are caused by movement of the Earth's tectonic plates, human activity can also produce earthquakes. Activities both above ground and below may change the stresses and strains on the crust, including building reservoirs, extracting resources such as coal or oil, and injecting fluids underground for waste disposal or fracking. Most of these earthquakes have small magnitudes. The 5.7 magnitude 2011 Oklahoma earthquake is thought to have been caused by disposing wastewater from oil production into injection wells, and studies point to the state's oil industry as the cause of other earthquakes in the past century. A Columbia University paper suggested that the 8.0 magnitude 2008 Sichuan earthquake was induced by loading from the Zipingpu Dam, though the link has not been conclusively proved.

Measuring and locating earthquakes

The instrumental scales used to describe the size of an earthquake began with the Richter magnitude scale in the 1930s. It is a relatively simple measurement of an event's amplitude, and its use has become minimal in the 21st century. Seismic waves travel through the Earth's interior and can be recorded by seismometers at great distances. The surface wave magnitude was developed in the 1950s as a means to measure remote earthquakes and to improve the accuracy for larger events. The moment magnitude scale not only measures the amplitude of the shock but also takes into account the seismic moment (total rupture area, average slip of the fault, and rigidity of the rock). The Japan Meteorological Agency seismic intensity scale, the Medvedev–Sponheuer–Karnik scale, and the Mercalli intensity scale are based on the observed effects and are related to the intensity of shaking.

Every tremor produces different types of seismic waves, which travel through rock with different velocities:

- Longitudinal P-waves (shock- or pressure waves)
- Transverse S-waves (both body waves)
- Surface waves – (Rayleigh and Love waves)

Propagation velocity of the seismic waves through solid rock ranges from approx. 3 km/s (1.9 mi/s) up to 13 km/s (8.1 mi/s), depending on the density and elasticity of the medium. In the Earth's interior, the shock- or P-waves travel much faster than the S-waves (approx. relation 1.7:1). The differences in travel time from the epicenter to the observatory are a measure of the distance and can be used to image both sources of quakes and structures within the Earth. Also, the depth of the hypocenter can be computed roughly.

In the upper crust, P-waves travel in the range 2–3 km (1.2–1.9 mi) per second (or lower) in soils and unconsolidated sediments, increasing to 3–6 km (1.9–3.7 mi) per second in solid rock. In the lower crust, they travel at about 6–7 km (3.7–4.3 mi) per second; the velocity increases within the deep mantle to about 13 km (8.1 mi) per second. The velocity of S-waves ranges from 2–3 km (1.2–1.9 mi) per second in light sediments and 4–5 km (2.5–3.1 mi) per second in the Earth's crust up to 7 km (4.3 mi) per second in the deep mantle. As a consequence, the first waves of a distant earthquake arrive at an observatory via the Earth's mantle.

On average, the kilometer distance to the earthquake is the number of seconds between the P- and S-wave times 8. Slight deviations are caused by inhomogeneities of subsurface structure. By such analyses of seismograms the Earth's core was located in 1913 by Beno Gutenberg.

S-waves and later arriving surface waves do most of the damage compared to P-waves. P-waves squeeze and expand material in the same direction they are traveling, whereas S-waves shake the ground up and down and back and forth.

Earthquakes are not only categorized by their magnitude but also by the place where they occur. The world is divided into 754 Flinn–Engdahl regions (F-E regions), which are based on political and geographical boundaries as well as seismic activity. More active zones are divided into smaller F-E regions whereas less active zones belong to larger F-E regions.

Standard reporting of earthquakes includes its magnitude, date and time of occurrence, geographic coordinates of its epicenter, depth of the epicenter, geographical region, distances to population centers, location uncertainty, a number of parameters that are included in USGS earthquake reports (number of stations reporting, number of observations, etc.), and a unique event ID.

Although relatively slow seismic waves have traditionally been used to detect earthquakes, scientists realized in 2016 that gravitational measurements could provide instantaneous detection of earthquakes, and confirmed this by analyzing gravitational records associated with the 2011 Tohoku-Oki ("Fukushima") earthquake.

SPATIAL SCALES – MICRO, MESO AND MACRO

We compare the relative importance of four dimensions for explaining the micro location of robberies: 1) the micro spatial scale of street segments; 2) the meso spatial scale surrounding the street segment; 3) the temporal pattern, and 4) the macro-scale of the surrounding 2.5 miles.

Mesoregion (geomorphology)

In some geomorphological taxonomies, a mesoregion is a natural region of intermediate size.

Mesoregions may be defined on the basis of morpholithogenic conditions and spatial connection. A mesoregion is used in the regionalization of the area, it is a physical and geographic division unit, it is part of a macroregion, and consists of microregions.

Spatial Scales

At global and continental scales, geomorphologists study major physiographic features like mountain belts, depositional basins, and great river systems like the Mississippi, Amazon, or Nile. At these scales, broad patterns in global climate and plate tectonics influence patterns of erosion and deposition that, in turn, influence the size and extent of mountains, plateaus, lowlands, coastal plains, and river basins. Such patterns also affect the distribution of soil types and the relative importance of glacial, fluvial, eolian (wind-driven), and coastal processes in shaping topography.

At a regional scale, distinct physiographic provinces reflect areas with broadly similar history, landforms, and landscape dynamics. Examples of such provinces include mountain belts like the Sierra Nevada or Appalachians as well as features like the Great Plains, California's Central Valley, and the Colorado Plateau in the American Southwest. Physiographic provinces are areas in which similar suites of geomorphological processes govern landscape formation and dynamics, and thus where one finds similar suites of landforms. A drainage basin is the land surface area drained by a given stream. Small streams come together to form larger rivers, so landscapes are naturally organized into smaller drainage basins nested within larger drainage basins.

Drainage basins range in size from a headwater catchment that collects water from a single mountain hillside to the Amazon River basin that drains more than half of South America. Drainage divides are topographic ridges that separate drainage basins.

Because streamflow is the predominant agent of erosion and sediment transport on land, drainage basins are the logical unit for analysis of many geomorphological processes. At the scale of individual valley segments, the tectonic or climatic setting of a region often defines areas where different types of processes and/or histories have led to development of distinct landforms and dynamics. The difference between U-shaped valleys carved by glaciers valleys and V-shaped valleys cut by streams is a classic example. Systematic downstream changes in stream valley morphology, from erosional headwater streams with narrow valleys that are confined between bedrock walls, to broad unconfined valleys of depositional lowlands. Distinct suites of valley segment types are diagnostic of specific physiographic provinces, and their character and distribution both vary regionally and generally reflect the history and processes of landscape evolution and shape ecosystem dynamics. At finer spatial scales, landscapes can be divided into distinct hillslopes, hollows, channels, floodplains, and estuaries. Hillslopes (including hilltops) are the undissected uplands between valleys. Hollows are unchanneled valleys that typically occur at the head of channels in soil-mantled terrain. Channels are zones of concentrated flow and sediment transport within well-defined banks, and floodplains are the flat valley bottoms along river valleys that are inundated during times of high discharge under the present climate. Estuaries are locations where streams enter coastal waters to arrive at their ultimate destination — sea level.

Temporal Scales

The scale of observations in both time and space strongly influence geomorphic interpretations. Matching the scale of observation to the scale of the question you seek to answer is critical for gathering meaningful data. It's easy to understand the importance of spatial scale. Measuring erosion on a single hillside in Kenya won't tell you much about the erosion rate of the African continent. It is more difficult to understand the influence of time on geomorphic data, observations, and interpretations.

Consider how the first estimates of continental erosion rates, which were made by measuring the concentration of suspended sediment exported by rivers over a few years, turned out to be wrong. While this is a reasonable approach if the measurements capture the variability of the system over time, in this case they did not. With only a few years of data, sediment yield results are likely to be biased by the short period of observation because most streams systems experience rare but massive floods that periodically transport immense amounts of sediment. How would one know that the few years in which the sampling occurred are representative of a meaningful, long-term average? The issue of how to integrate the influence of large and small events in shaping larger landforms remains central to modern process geomorphology.

Topography evolves over periods of time that range from the millions of years that are required to erode away mountain belts, to the few seconds, minutes, or days it takes for a landslide, flood, or earthquake-driven fault displacement to disrupt the land surface. Climate cycles influence topography over millennia as glaciers advance, retreat, and scour out alpine valleys. Likewise, river profiles adjust to the sea level changes that accompany glaciations. Landscape responses to large-scale disturbances like hurricanes and volcanic eruptions are often evident for centuries, and it can take decades for landslide scars to revegetate and river channels to process the sediment shed from slopes during large storms. River flow exhibits annual and seasonal variability that controls the timing of sediment movement and structure of stream ecosystems. Because of the disproportionate influence of infrequent extreme events like storms, landslides, and floods, rates of processes measured over short time spans may not adequately characterize average rates over longer time scales. Not surprisingly, geomorphologists deal with a wide variety of measurements over different time scales, from rates directly measured in the field, to indirect measurements of long term erosion rates inferred from isotopic analyses, and erosion rates constrained by sedimentary volumes preserved in depositional basins. The key, of course, is to employ analyses relevant to the time scale of interest.

SPATIAL ANALYSIS

Spatial analysis or spatial statistics includes any of the formal techniques which study entities using their topological, geometric, or geographic properties. The phrase properly refers to a variety of techniques, many still in their early development, using different analytic approaches and applied in fields as diverse as astronomy, with its studies of the placement of galaxies in the cosmos, to chip fabrication engineering, with its use of 'place and route' algorithms to build complex wiring structures. The phrase is often used in a more restricted sense to describe techniques applied to structures at the human scale, most notably in the analysis of geographic data. The phrase is even sometimes used to refer to a specific technique in a single area of research, for example, to describe geostatistics.

Complex issues arise in spatial analysis, many of which are neither clearly defined nor completely resolved, but form the basis for current research. The most fundamental of these is the problem of defining the spatial location of the entities being studied. For example, a study on human health could describe the spatial position of humans with a point placed where they live, or with a point located where they work, or by using a line to describe their weekly trips; each choice has dramatic effects on the techniques which can be used for the analysis and on the conclusions which can be obtained. Other issues in spatial analysis include the limitations of mathematical knowledge, the assumptions required by existing statistical techniques, and problems in computer based calculations.

Classification of the techniques of spatial analysis is difficult because of the large number of different fields of research involved, the different fundamental approaches which can be chosen, and the many forms the data can take.

The History of Spatial Analysis

Spatial analysis can perhaps be considered to have arisen with the early attempts at cartography and surveying but many fields have contributed to its rise in modern form. Biology contributed through botanical studies of global plant distributions and local plant locations, ethological studies of animal movement, landscape ecological studies of vegetation blocks, ecological studies of spatial population dynamics, and the study of biogeography.

Epidemiology contributed with early work on disease mapping, notably John Snow's work mapping an outbreak of cholera, with research on mapping the spread of disease and with locational studies for health care delivery. Statistics has contributed greatly through work in spatial statistics.

Economics has contributed notably through spatial econometrics. Geographic information system is currently a major contributor due to the importance of geographic software in the modern analytic toolbox. Remote sensing has contributed extensively in morphometric and clustering analysis. Computer science has contributed extensively through the study of algorithms, notably in computational geometry. Mathematics continues to provide the fundamental tools for analysis and to reveal the complexity of the spatial realm, for example, with recent work on fractals and scale invariance. Scientific modelling provides a useful framework for new approaches.

Fundamental Issues in Spatial Analysis

Spatial analysis confronts many fundamental issues in the definition of its objects of study, in the construction of the analytic operations to be used, in the use of computers for analysis, in the limitations and particularities of the analyses which are known, and in the presentation of analytic results. Many of these issues are active subjects of modern research.

Common errors often arise in spatial analysis, some due to the mathematics of space, some due to the particular ways data are presented spatially, some due to the tools which are available. Census data, because it protects individual privacy by aggregating data into local units, raises a number of statistical issues. The fractal nature of coastline makes precise measurements of its length difficult if not impossible. A computer software fitting straight lines to the curve of a coastline, can easily calculate the lengths of the lines which it defines. However these straight lines may have no inherent meaning in the real world, as was shown for the coastline of Britain.

These problems represent one of the greatest dangers in spatial analysis because of the inherent power of maps as media of presentation. When results are presented as maps, the presentation combines the spatial data which is generally very accurate with analytic results which may be grossly inaccurate. Some of these issues are discussed at length in the book *How to Lie with Maps*

Spatial Characterization

The definition of the spatial presence of an entity constrains the possible analysis which can be applied to that entity and influences the final conclusions that can be reached. While this property is fundamentally true of all analysis, it is particularly important in spatial analysis because the tools to define and study entities favour specific characterizations of the entities being studied. Statistical techniques favour the spatial definition of objects as points because there are very few statistical techniques which operate directly on line, area, or volume elements. Computer tools favour the spatial definition of objects as homogeneous and separate elements because of the limited number of database elements and computational structures available, and the ease with which these primitive structures can be created.

Spatial Dependency or Auto-correlation

Spatial dependency is the co-variation of properties within geographic space: characteristics at proximal locations appear to be correlated, either positively or negatively. Spatial dependency leads to the spatial auto correlation problem in statistics since, like temporal auto correlation, this violates standard statistical techniques that assume independence among observations. For example, regression analyses that do not compensate for spatial dependency can have unstable parameter estimates and yield unreliable significance tests. Spatial regression models capture these relationships and do not suffer from these weaknesses. It is also appropriate to view spatial dependency as a source of information rather than something to be corrected.

Locational effects also manifest as spatial heterogeneity, or the apparent variation in a process with respect to location in geographic space. Unless a space is uniform and boundless, every location will have some degree of uniqueness relative to the other locations. This affects the spatial dependency relations and therefore the spatial process. Spatial heterogeneity means that overall parameters estimated for the entire system may not adequately describe the process at any given location.

Scaling

Spatial measurement scale is a persistent issue in spatial analysis; more detail is available at the modifiable areal unit problem (MAUP) topic entry. Landscape ecologists developed a series of scale invariant metrics for aspects of ecology that are fractal in nature. In more general terms, no scale independent method of analysis is widely agreed upon for spatial statistics.

Sampling

Spatial sampling involves determining a limited number of locations in geographic space for faithfully measuring phenomena that are subject to dependency and heterogeneity. Dependency suggests that since one location can predict the value of another location, we do not need observations in both places. But heterogeneity suggests that this relation can change across space, and therefore we cannot trust an observed degree of dependency beyond a region that may be small. Basic spatial sampling schemes include random, clustered and systematic. These basic schemes can be applied at multiple levels in a designated spatial hierarchy (e.g., urban area, city, neighbourhood). It is also possible to exploit ancillary data, for example, using property values as a guide in a spatial sampling scheme to measure educational attainment and income. Spatial models such as auto correlation statistics, regression and interpolation can also dictate sample design.

Common Errors in Spatial Analysis

The fundamental issues in spatial analysis lead to numerous problems in analysis including bias, distortion and outright errors in the conclusions reached. These issues are often interlinked but various attempts have been made to separate out particular issues from each other.

Length

In a paper by Benoit Mandelbrot on the coastline of Britain it was shown that it is inherently nonsensical to discuss certain spatial concepts despite an inherent presumption of the validity of the concept. Lengths in ecology depend directly on the scale at which they are measured and experienced. So while surveyors commonly measure the length of a river, this length only has meaning in the context of the relevance of the measuring technique to the question under study.

Locational Fallacy

The locational fallacy refers to error due to the particular spatial characterization chosen for the elements of study, in particular choice of placement for the spatial presence of the element.

Spatial characterizations may be simplistic or even wrong. Studies of humans often reduce the spatial existence of humans to a single point, for instance their home address. This can easily lead to poor analysis, for example, when considering disease transmission which can happen at work or at school and therefore far from the home.

The spatial characterization may implicitly limit the subject of study. For example, the spatial analysis of crime data has recently become popular but these studies can only describe the particular kinds of crime which can be described spatially.

This leads to many maps of assault but not to any maps of embezzlement with political consequences in the conceptualization of crime and the design of policies to address the issue.

Atomic Fallacy

This describes errors due to treating elements as separate 'atoms' outside of their spatial context.

Ecological Fallacy

The ecological fallacy describes errors due to performing analyses on aggregate data when trying to reach conclusions on the individual units.

Errors occur in part from spatial aggregation. For example a pixel represents the average surface temperatures within an area. Ecological fallacy would be to assume that all points within the area have the same temperature. This topic is closely related to the modifiable areal unit problem.

CLIMATIC GEOMORPHOLOGY AND TECTONIC GEOMORPHOLOGY

Climatic geomorphology

During the age of New Imperialism in the late 19th century European explorers and scientists traveled across the globe bringing descriptions of landscapes and landforms. As geographical knowledge increased over time these observations were systematized in a search for regional patterns. Climate emerged thus as prime factor for explaining landform distribution at a grand scale. The rise of climatic geomorphology was foreshadowed by the work of Wladimir Köppen, Vasily Dokuchaev and Andreas Schimper. William Morris Davis, the leading geomorphologist of his time, recognized the role of climate by complementing his "normal" temperate climate cycle of erosion with arid and glacial ones. Nevertheless, interest in climatic geomorphology was also a reaction *against* Davisian geomorphology that was by the mid-20th century considered both un-innovative and dubious. Early climatic geomorphology developed primarily in continental Europe while in the English-speaking world the tendency was not explicit until L.C. Peltier's 1950 publication on a periglacial cycle of erosion.

The criticism by Stoddart proved "devastating" sparking a decline in the popularity of climatic geomorphology in the late 20th century. Stoddart criticized climatic geomorphology for applying supposedly "trivial" methodologies in establishing landform differences between morphoclimatic zones, being linked to Davisian geomorphology and by allegedly neglecting the fact that physical laws governing processes are the same across the globe. In addition some conceptions of climatic geomorphology, like that which holds that chemical weathering is more rapid in tropical climates than in cold climates proved to not be straightforwardly true.

Global Tectonic and Climatic Systems

Tectonic geomorphology involves the interactions among landforms, landscapes, and tectonics. Tectonics, the branch of geology dealing with regional structures and deformation features, occupies a central role in the Earth sciences. The discipline has achieved great importance through the unifying role of the plate tectonic model in explaining the large-scale surface features of our planet.

Major advances in tectonic geomorphology have been made in the last decade, mainly because of an increased ability to evaluate the time factor in landscape development. Thus, through the use of geochemical means of dating and computerized models of landform change, it is now possible to evaluate differential rates of uplift.

It is also possible to determine the magnitude and frequency of displacements along faults. Perhaps the most important new development for climatic and climatogenetic geomorphology is the use of analytical models to characterize the global interactions of the land surface, atmosphere, and oceans. Most interesting are the general circulation models that simulate global atmospheric processes. For climatic geomorphology, these models have shown profound feedback relationships operating between soil moisture and precipitation, carbon dioxide and climate and anthropogenic changes in the albedo and climate.

An example of an especially useful climate-genetic reference point is provided by the CLIMAP reconstruction of the 18000 years B.P. global climate from a compilation of the Earth´s ocean surface temperatures. A preliminary extension of the analysis to continental areas revealed numerous problems of local variability.

Such comparisons between general atmospheric conditions, as modeled by GCMs, and paleogeomorphic reconstructions hold great promise for understanding large-scale climate/landscape interactions. The goal here is to generate a self-enhancing spiral of understanding, with models pointing to key geomorphic questions and geomorphic data refining the models.

Ancient Landscapes

At the end of the l9th century, geomorphology achieved an important theoretical synthesis through the work of William Morris Davis. Davis conceived a marvelous deductive scheme of landscape development by the action of exogenetic processes acting on the basic materials and structures to produce a progressive evolution of landscape stages through time.

Unfortunately, this theoretical framework was somewhat abused by those who employed it solely for landscape description and classification. By the middle of the 20th century, evolutionary geomorphology fell from favour among Earth scientists, who focused their primary efforts on the study of various geomorphic processes.

An unfortunate by-product of the controversy over Davisian geomorp-hology was the general abandonment, especially in Britain and the United States, of studies that concerned ancient landscapes and landforms. The development of radiometric dating has rekindled interest in this topic by identifying the antiquity of landscapes.

For example, Young has shown that upland surfaces in southeastern Australia originated as early as Mesozoic and Early Tertiary, with the landscape assuming its approximate present-day form by the Miocene. Twidale et al. also identified Mesozoic landscapes in South Australia. Ollier in an analysis of ancient landscapes in Australia, concluded that conventional approaches to geomorphic change suffer from inadequate appreciation of broad scales of time and space. He proposed an evolutionary approach to geomorphic features, such as that applied by geologists to tectonic features. Low-relief plains cutting across varied rocks and structures are common features on the Earth. Such surfaces have long been of interest to geomorphologists, and many scientific controversies have arisen over their explanation.

The genetic implications are contained in the many names for the surfaces; peneplains, pediplains, panplains, etchplains, exhumed plains, and paleoplains. To avoid the problems inherent in these controversies, it is perhaps best to simply call these features planation surfaces.

Table. Widespread Planation Surfaces

Name	Age	Comments
Gondwana	Jurassic	Related to Pangea and its breakup.
Post-Gondwana or Kretacic	Early to Mid-Cretaceous	Related to Pangea and its breakup.
African or Moorland	Late cretaceous to Early Cenozoic	Extensive surface created by stripping weathered material from older surfaces.
Post-African or Rolling	Miocene	Undulating surface developed above younger valleys
Widespread	Pliocene	Global surface common near coastal areas.
Youngest	Quaternary	Latest valley formation.

A spectacular planation surface that bevels sandstone cuestas in the fold belt of the Amadeus Basin in central Australia. An analysis of the regional geomorphology of this area is presented in Plate I-4.

The question of regional planation surfaces obviously awaits a modern global analysis.

Perhaps the classic syntheses of King can be re-evaluated by the use of the new techniques described in this volume. It would be refreshing to accomplish such a study with automated data collection procedures, free of the raging controversy that so hampered the highly personal interpretive studies of the past.

The identification and dating of various planation surfaces have become important components of tectonic geomorphic analysis. Deformation of such surfaces by faulting, folding, or broad warping can be calibrated by the displacement of a planation surface from its original attitude.

Climatic Geomorphology Approach

In Europe, outside Britain and France, Geomorphology progressed more or less without to reference to Davis erosion cycle or to denudation chronology.

Their criticism of these historical approaches reduced to its simplest was that different climates produce different processes which in turn produced different landforms.

In place of these historical approaches, they adopted an alternative theoretical approach which has been called Climatic Geomorphology. The theme of the approach is that distinctive climates possess distinctive assemblages of processes which result in different assemblages of landforms.

According to this approach, every phenomenon or process whose global extension is more or less comfortable to latitude is termed zonal. The end product of the climatic approach is the identification of a number of so called morphoclimatic zones of the earth, each with its distinctive climate processes and landforms.

Thus, the identification of regions where climate may determine the dominant geomorphic processes and therefore significantly influence landform production is one of the goals of climatic geomorphologists. However, the development of climatic geomorphology has been traced to Von Richthjofen who at the close of the 19th century was developing his ideas of climatic geomorphology.

This early start was continued much later by J. Budel (also a German) followed by such worker as Birot, Tricart and Cailleux (all French) and Derbyshire, a Briton. The climatic geomorphology has been gaining strength not only in Europe, but also in North America since 1950 and it is still applied in elucidating the origin and evolution of regional landform assemblages world-wide.

Dynamic of Plate Tectonics

Plate tectonics, theory dealing with the dynamics of Earth's outer shell—the lithosphere—that revolutionized Earth sciences by providing a uniform context for understanding mountain-building processes, volcanoes, and earthquakes as well as the evolution of Earth's surface and reconstructing its past continents and oceans.

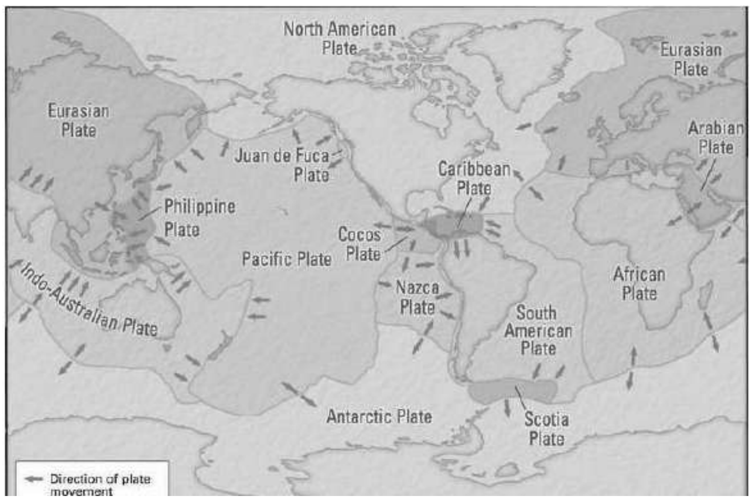

Fig. Earth's tectonic plates

The concept of plate tectonics was formulated in the 1960s. According to the theory, Earth has a rigid outer layer, known as the lithosphere, which is typically about 100 km (60 miles) thick and overlies a plastic (moldable, partially molten) layer called the asthenosphere. The lithosphere is broken up into seven very large continental- and ocean-sized plates, six or seven medium-sized regional plates, and several small ones. These plates move relative to each other, typically at rates of 5 to 10 cm (2 to 4 inches) per year, and interact along their boundaries, where they converge, diverge, or slip past one another.

Such interactions are thought to be responsible for most of Earth's seismic and volcanic activity, although earthquakes and volcanoes can occur in plate interiors. Plate motions cause mountains to rise where plates push together, or converge, and continents to fracture and oceans to form where plates pull apart, or diverge. The continents are embedded in the plates and drift passively with them, which over millions of years results in significant changes in Earth's geography.

The theory of plate tectonics is based on a broad synthesis of geologic and geophysical data. It is now almost universally accepted, and its adoption represents a true scientific revolution, analogous in its consequences to quantum mechanics in physics or the discovery of the genetic code in biology.

Incorporating the much older idea of continental drift, as well as the concept of seafloor spreading, the theory of plate tectonics has provided an overarching framework in which to describe the past geography of continents and oceans, the processes controlling creation and destruction of landforms, and the evolution of Earth's crust, atmosphere, biosphere, hydrosphere, and climates.

During the late 20th and early 21st centuries, it became apparent that plate-tectonic processes profoundly influence the composition of Earth's atmosphere and oceans, serve as a prime cause of long-term climate change, and make significant contributions to the chemical and physical environment in which life evolves.

Principles Of Plate Tectonics

In essence, plate-tectonic theory is elegantly simple. Earth's surface layer, 50 to 100 km (30 to 60 miles) thick, is rigid and is composed of a set of large and small plates. Together, these plates constitute the lithosphere, from the Greek *lithos*, meaning "rock."

The lithosphere rests on and slides over an underlying partially molten (and thus weaker but generally denser) layer of plastic partially molten rock known as the asthenosphere, from the Greek *asthenos*, meaning "weak." Plate movement is possible because the lithosphere-asthenosphere boundary is a zone of detachment. As the lithospheric plates move across Earth's surface, driven by forces as yet not fully understood, they interact along their boundaries, diverging, converging, or slipping past each other. While the interiors of the plates are presumed to remain essentially undeformed, plate boundaries are the sites of many of the principal processes that shape the terrestrial surface, including earthquakes, volcanism, and orogeny (that is, formation of mountain ranges).

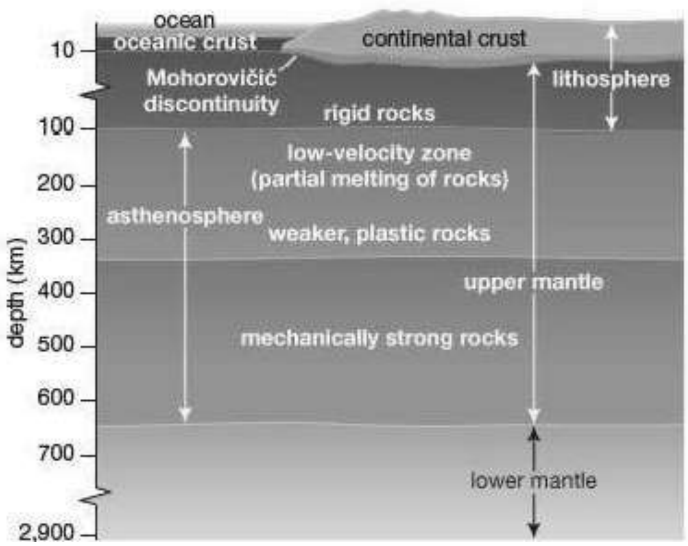

Fig. Earth's lithosphere and upper mantle

A cross section of Earth's outer layers, from the crust through the lower mantle. The process of plate tectonics may be driven by convection in Earth's mantle, the pull of heavy old pieces of crust into the mantle, or some combination of both.

Earth's layers

Knowledge of Earth's interior is derived primarily from analysis of the seismic waves that propagate through Earth as a result of earthquakes. Depending on the material they travel through, the waves may either speed up, slow down, bend, or even stop if they cannot penetrate the material they encounter.

Three-dimensional diagram showing crustal generation and destruction according to the theory of plate tectonics; included are the three kinds of plate boundaries—divergent, convergent (or collision), and strike-slip (or transform).

Collectively, these studies show that Earth can be internally divided into layers on the basis of either gradual or abrupt variations in chemical and physical properties. Chemically, Earth can be divided into three layers. A relatively thin crust, which typically varies from a few kilometres to 40 km (about 25 miles) in thickness, sits on top of the mantle. (In some places, Earth's crust may be up to 70 km [40 miles] thick.)

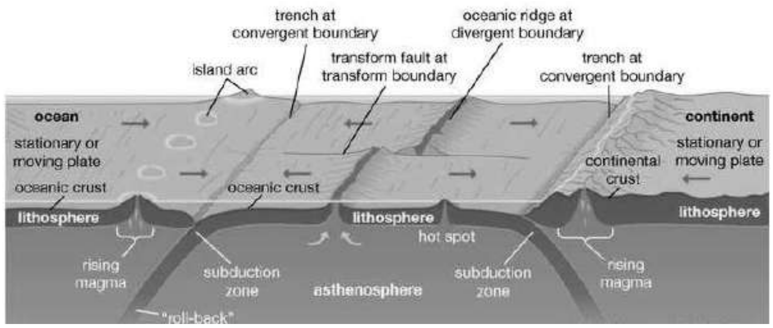

Fig. Crustal generation and destruction

The mantle is much thicker than the crust; it contains 83 percent of Earth's volume and continues to a depth of 2,900 km (1,800 miles). Beneath the mantle is the core, which extends to the centre of Earth, some 6,370 km (nearly 4,000 miles) below the surface. Geologists maintain that the core is made up primarily of metallic iron accompanied by smaller amounts of nickel, cobalt, and lighter elements, such as carbon and sulfur.

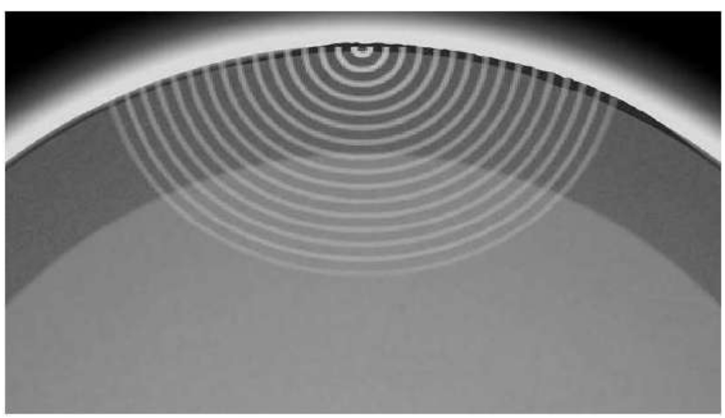

Fig. Discern between body and surface waves, primary and secondary waves, and Love and Rayleigh waves

The shifting rock in an earthquake causes vibrations called seismic waves that travel within Earth or along its surface. The four main types of seismic waves are *P* waves, *S* waves, Love waves, and Rayleigh waves.

There are two types of crust, continental and oceanic, which differ in their composition and thickness. The distribution of these crustal types broadly coincides with the division into continents and ocean basins, although continental shelves, which are submerged, are underlain by continental crust. The continents have a crust that is broadly granitic in composition and, with a density of about 2.7 grams per cubic cm (0.098 pound per cubic inch), is somewhat lighter than oceanic crust, which is basaltic (i.e., richer in iron and magnesium than granite) in composition and has a density of about 2.9 to 3 grams per cubic cm (0.1 to 0.11 pound per cubic inch). Continental crust is typically 40 km (25 miles) thick, while oceanic crust is much thinner, averaging about 6 km (4 miles) in thickness. These crustal rocks both sit on top of the mantle, which is ultramafic in composition (i.e., very rich in magnesium and iron-bearing silicate minerals). The boundary between the crust (continental or oceanic) and the underlying mantle is known as the Mohorovièiæ discontinuity (also called Moho), which is named for its discoverer, Croatian seismologist Andrija Mohorovièiæ. The Moho is clearly defined by seismic studies, which detect an acceleration in seismic waves as they pass from the crust into the denser mantle. The boundary between the mantle and the core is also clearly defined by seismic studies, which suggest that the outer part of the core is a liquid.

The effect of the different densities of lithospheric rock can be seen in the different average elevations of continental and oceanic crust. The less-dense continental crust has greater buoyancy, causing it to float much higher in the mantle. Its average elevation above sea level is 840 metres (2,750 feet), while the average depth of oceanic crust is 3,790 metres (12,400 feet). This density difference creates two principal levels of Earth's surface.

The lithosphere itself includes all the crust as well as the upper part of the mantle (i.e., the region directly beneath the Moho), which is also rigid. However, as temperatures increase with depth, the heat causes mantle rocks to lose their rigidity. This process begins at about 100 km (60 miles) below the surface. This change occurs within the mantle and defines the base of the lithosphere and the top of the asthenosphere. This upper portion of the mantle, which is known as the lithospheric mantle, has an average density of about 3.3 grams per cubic cm (0.12 pound per cubic inch). The asthenosphere, which sits directly below the lithospheric mantle, is thought to be slightly denser at 3.4–4.4 grams per cubic cm (0.12–0.16 pound per cubic inch).

In contrast, the rocks in the asthenosphere are weaker, because they are close to their melting temperatures. As a result, seismic waves slow as they enter the asthenosphere. With increasing depth, however, the greater pressure from the weight of the rocks above causes the mantle to become gradually stronger, and seismic waves increase in velocity, a defining characteristic of the lower mantle. The lower mantle is more or less solid, but the region is also very hot, and thus the rocks can flow very slowly (a process known as creep).

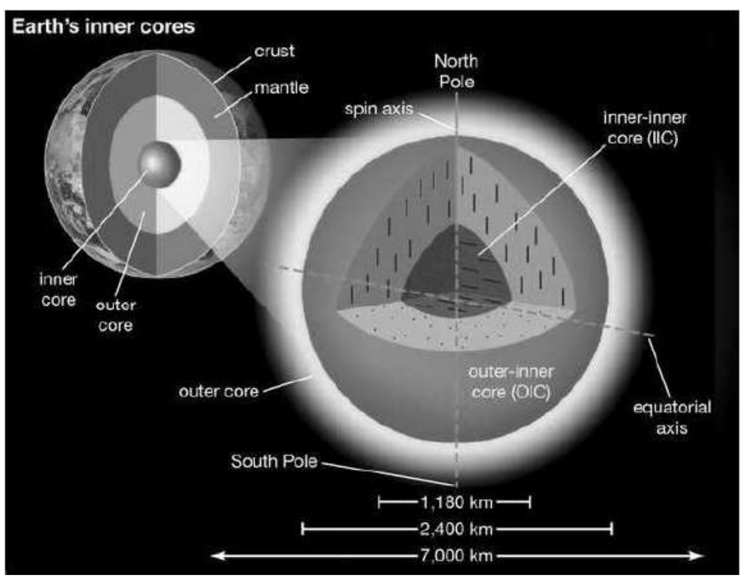

Fig. Earth's core

During the late 20th and early 21st centuries, scientific understanding of the deep mantle was greatly enhanced by high-resolution seismological studies combined with numerical modeling and laboratory experiments that mimicked conditions near the core-mantle boundary. Collectively, these studies revealed that the deep mantle is highly heterogeneous and that the layer may play a fundamental role in driving Earth's plates.

At a depth of about 2,900 km (1,800 miles), the lower mantle gives way to Earth's outer core, which is made up of a liquid rich in iron and nickel. At a depth of about 5,100 km (3,200 miles), the outer core transitions to the inner core. Although it has a higher temperature than the outer core, the inner core is solid because of the tremendous pressures that exist near Earth's centre. Earth's inner core is divided into the outer-inner core (OIC) and the inner-inner core (IIC), which differ from one another with respect to the polarity of their iron crystals. The polarity of the iron crystals of the OIC is oriented in a north-south direction, whereas that of the IIC is oriented east-west.

The internal layers of Earth's core, including its two inner cores.

Plate boundaries

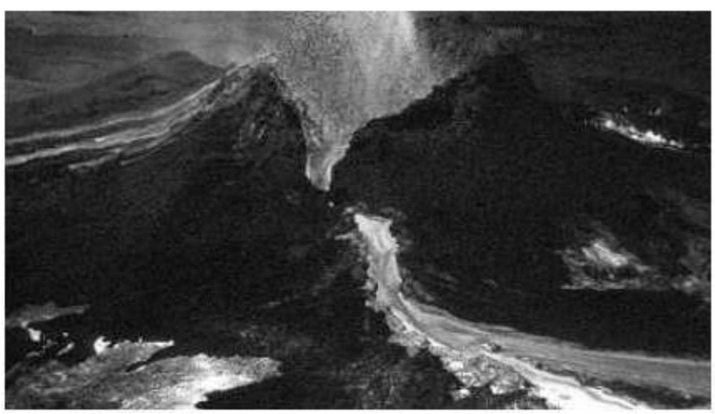

Fig. Examine how the theory of plate tectonics explains volcanic activity, earthquakes, and mountains

Lithospheric plates are much thicker than oceanic or continental crust. Their boundaries do not usually coincide with those between oceans and continents, and their behaviour is only partly influenced by whether they carry oceans, continents, or both.

The Pacific Plate, for example, is entirely oceanic, whereas the North American Plate is capped by continental crust in the west (the North American continent) and by oceanic crust in the east and extends under the Atlantic Ocean as far as the Mid-Atlantic Ridge.

In a simplified example of plate motion shown in the figure, movement of plate A to the left relative to plates B and C results in several types of simultaneous interactions along the plate boundaries. At the rear, plates A and B move apart, or diverge, resulting in extension and the formation of a divergent margin. At the front, plates A and B overlap, or converge, resulting in compression and the formation of a convergent margin.

Along the sides, the plates slide past one another, a process called shear. As these zones of shear link other plate boundaries to one another, they are called transform faults.

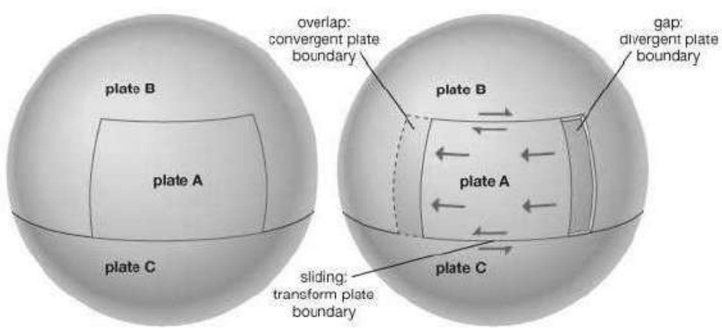

Fig. Plate movement

Theoretical diagram showing the effects of an advancing tectonic plate on other adjacent, but stationary, tectonic plates. At the advancing edge of plate A, the overlap with plate B creates a convergent boundary. In contrast, the gap left behind the trailing edge of plate A forms a divergent boundary with plate B. As plate

A slides past portions of both plate B and plate C, transform boundaries develop.

Divergent margins

As plates move apart at a divergent plate boundary, the release of pressure produces partial melting of the underlying mantle. This molten material, known as magma, is basaltic in composition and is buoyant. As a result, it wells up from below and cools close to the surface to generate new crust. Because new crust is formed, divergent margins are also called constructive margins.

Continental rifting

Upwelling of magma causes the overlying lithosphere to uplift and stretch. (Whether magmatism [the formation of igneous rock from magma] initiates the rifting or whether rifting decompresses the mantle and initiates magmatism is a matter of significant debate.)

Fig. Rift valley in Thingvellir National Park

If the diverging plates are capped by continental crust, fractures develop that are invaded by the ascending magma, prying the continents farther apart. Settling of the continental blocks creates a rift valley, such as the present-day East African Rift Valley. As the rift continues to widen, the continental crust becomes progressively thinner until separation of the plates is achieved and a new ocean is created. The ascending partial melt cools and crystallizes to form new crust. Because the partial melt is basaltic in composition, the new crust is oceanic, and an ocean ridge develops along the site of the former continental rift. Consequently, diverging plate boundaries, even if they originate within continents, eventually come to lie in ocean basins of their own making.

The Thingvellir fracture zone at Thingvellir National Park in southwestern Iceland is an example of a rift valley. The Thingvellir fracture lies in the Mid-Atlantic Ridge, which extends through the centre of Iceland.

Seafloor spreading

As upwelling of magma continues, the plates continue to diverge, a process known as seafloor spreading. Samples collected from the ocean floor show that the age of oceanic crust increases with distance from the spreading centre—important evidence in favour of this process.

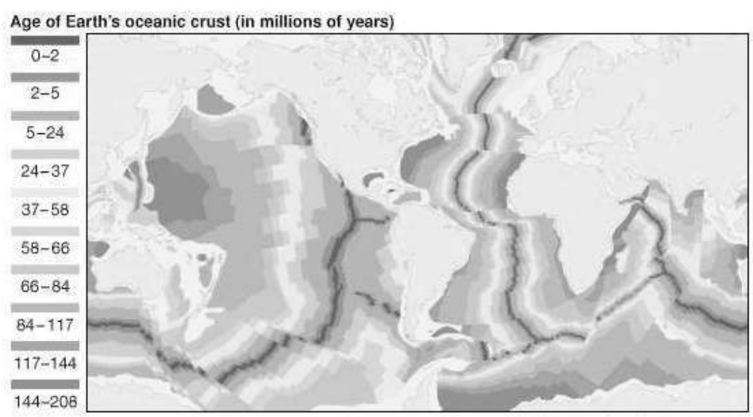

Fig. Age of Earth's oceanic crust

These age data also allow the rate of seafloor spreading to be determined, and they show that rates vary from about 0.1 cm (0.04 inch) per year to 17 cm (6.7 inches) per year. Seafloor-spreading rates are much more rapid in the Pacific Ocean than in the Atlantic and Indian oceans. At spreading rates of about 15 cm (6 inches) per year, the entire crust beneath the Pacific Ocean (about 15,000 km [9,300 miles] wide) could be produced in 100 million years.

The age of Earth's oceanic crust can be presented to show the pattern of seafloor spreading at the global scale.

Divergence and creation of oceanic crust are accompanied by much volcanic activity and by many shallow earthquakes as the crust repeatedly rifts, heals, and rifts again. Brittle earthquake-prone rocks occur only in the shallow crust. Deep earthquakes, in contrast, occur less frequently, due to the high heat flow in the mantle rock. These regions of oceanic crust are swollen with heat and so are elevated by 2 to 3 km (1.2 to 1.9 miles) above the surrounding seafloor. The elevated topography results in a feedback scenario in which the resulting gravitational force pushes the crust apart, allowing new magma to well up from below, which in turn sustains the elevated topography. Its summits are typically 1 to 5 km (0.6 to 3.1 miles) below the ocean surface. On a global scale, these ridges form an interconnected system of undersea "mountains" that are about 65,000 km (40,000 miles) in length and are called oceanic ridges.

Convergent margins

Given that Earth is constant in volume, the continuous formation of Earth's new crust produces an excess that must be balanced by destruction of crust elsewhere. This is accomplished at convergent plate boundaries, also known as destructive plate boundaries, where one plate descends at an angle—that is, is subducted—beneath the other.

Because oceanic crust cools as it ages, it eventually becomes denser than the underlying asthenosphere, and so it has a tendency to subduct, or dive under, adjacent continental plates or younger sections of oceanic crust. The life span of the oceanic crust is prolonged by its rigidity, but eventually this resistance is overcome. Experiments show that the subducted oceanic lithosphere is denser than the surrounding mantle to a depth of at least 600 km (about 400 miles).

The mechanisms responsible for initiating subduction zones are controversial. During the late 20th and early 21st centuries, evidence emerged supporting the notion that subduction zones preferentially initiate along preexisting fractures (such as transform faults) in the oceanic crust. Irrespective of the exact mechanism, the geologic record indicates that the resistance to subduction is overcome eventually.

Where two oceanic plates meet, the older, denser plate is preferentially subducted beneath the younger, warmer one. Where one of the plate margins is oceanic and the other is continental, the greater buoyancy of continental crust prevents it from sinking, and the oceanic plate is preferentially subducted. Continents are preferentially preserved in this manner relative to oceanic crust, which is continuously recycled into the mantle. This explains why ocean floor rocks are generally less than 200 million years old whereas the oldest continental rocks are more than 4 billion years old. Before the middle of the 20th century, most geoscientists maintained that continental crust was too buoyant to be subducted. However, it later became clear that slivers of continental crust adjacent to the deep-sea trench, as well as sediments deposited in the trench, may be dragged down the subduction zone. The recycling of this material is detected in the chemistry of volcanoes that erupt above the subduction zone.

Subduction zones

The subduction process involves the descent into the mantle of a slab of cold hydrated oceanic lithosphere about 100 km (60 miles) thick that carries a relatively thin cap of oceanic sediments. The path of descent is defined by numerous earthquakes along a plane that is typically inclined between 30° and 60° into the mantle and is called the Wadati-Benioff zone, for Japanese seismologist Kiyoo Wadati and American seismologist Hugo Benioff, who pioneered its study. Between 10 and 20 percent of the subduction zones that dominate the circum-Pacific ocean basin are subhorizontal (that is, they subduct at angles between 0° and 20°). The factors that govern the dip of the subduction zone are not fully understood, but they probably include the age and thickness of the subducting oceanic lithosphere and the rate of plate convergence.

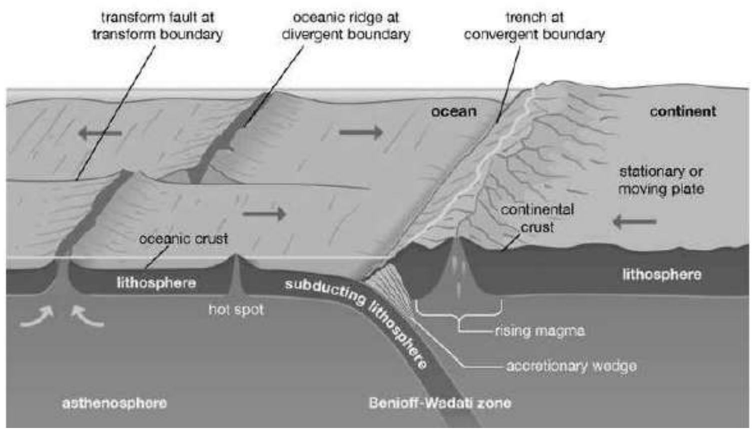

Fig. Subducting tectonic plate

A subducting plate's path (called the Benioff-Wadati [or Wadati-Benioff] zone) is defined by numerous earthquakes along a plane that is typically inclined between 30° and 60° into the mantle.

Most, but not all, earthquakes in this planar dipping zone result from compression, and the seismic activity extends 300 to 700 km (200 to 400 miles) below the surface, implying that the subducted crust retains some rigidity to this depth. At greater depths the subducted plate is partially recycled into the mantle.

The site of subduction is marked by a deep trench, between 5 and 11 km (3 and 7 miles) deep, that is produced by frictional drag between the plates as the descending plate bends before it subducts.

The overriding plate scrapes sediments and elevated portions of ocean floor off the upper crust of the lower plate, creating a zone of highly deformed rocks within the trench that becomes attached, or accreted, to the overriding plate. This chaotic mixture is known as an accretionary wedge.

The rocks in the subduction zone experience high pressures but relatively low temperatures, an effect of the descent of the cold oceanic slab. Under these conditions the rocks recrystallize, or metamorphose, to form a suite of rocks known as blueschists, named for the diagnostic blue mineral called glaucophane, which is stable only at the high pressures and low temperatures found in subduction zones. At deeper levels in the subduction zone (that is, greater than 30–35 km [about 19–22 miles]), eclogites, which consist of high-pressure minerals such as red garnet (pyrope) and omphacite (pyroxene), form. The formation of eclogite from blueschist is accompanied by a significant increase in density and has been recognized as an important additional factor that facilitates the subduction process.

Island arcs

When the downward-moving slab reaches a depth of about 100 km (60 miles), it gets sufficiently warm to drive off its most volatile components, thereby stimulating partial melting of mantle in the plate above the subduction zone (known as the mantle wedge). Melting in the mantle wedge produces magma, which is predominantly basaltic in composition. This magma rises to the surface and gives birth to a line of volcanoes in the overriding plate, known as a volcanic arc, typically a few hundred kilometres behind the oceanic trench. The distance between the trench and the arc, known as the arc-trench gap, depends on the angle of subduction. Steeper subduction zones have relatively narrow arc-trench gaps. A basin may form within this region, known as a fore-arc basin, and may be filled with sediments derived from the volcanic arc or with remains of oceanic crust.

If both plates are oceanic, as in the western Pacific Ocean, the volcanoes form a curved line of islands, known as an island arc, that is parallel to the trench, as in the case of the Mariana Islands and the adjacent Mariana Trench. If one plate is continental, the volcanoes form inland, as they do in the Andes of western South America. Though the process of magma generation is similar, the ascending magma may change its composition as it rises through the thick lid of continental crust, or it may provide sufficient heat to melt the crust. In either case, the composition of the volcanic mountains formed tends to be more silicon-rich and iron- and magnesium-poor relative to the volcanic rocks produced by ocean-ocean convergence.

Back-arc basins

Where both converging plates are oceanic, the margin of the older oceanic crust will be subducted because older oceanic crust is colder and therefore more dense. As the dense slab collapses into the asthenosphere, however, it also may "roll back" oceanward and cause extension in the overlying plate. This results in a process known as back-arc spreading, in which a basin opens up behind the island arc. The crust behind the arc becomes progressively thinner, and the decompression of the underlying mantle causes the crust to melt, initiating seafloor-spreading processes, such as melting and the production of basalt; these processes are similar to those that occur at ocean ridges. The geochemistry of the basalts produced at back-arc basins superficially resembles that of basalts produced at ocean ridges, but subtle trace element analyses can detect the influence of a nearby subducted slab.

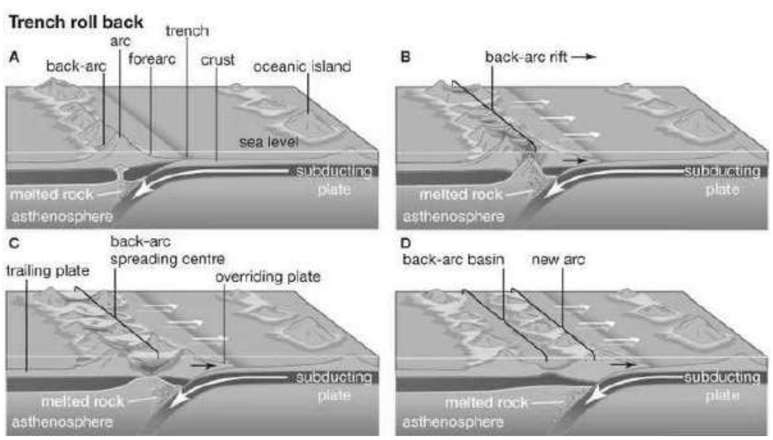

Fig. Back-arc basin

The trench "roll back" process of back-arc basin formation.

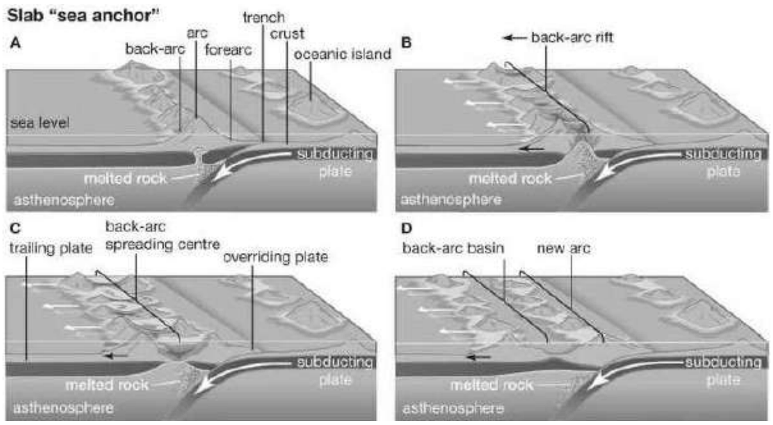

Fig. Sea anchor process in back-arc basin formation

The slab "sea anchor" process of back-arc basin formation.

This style of subduction predominates in the western Pacific Ocean, in which a number of back-arc basins separate several island arcs from Asia. Examples include the Mariana Islands, the Kuril Islands, and the main islands of Japan. However, if the rate of convergence increases or if anomalously thick oceanic crust (possibly caused by rising mantle plume activity) is conveyed into the subduction zone, the slab may flatten. Such flattening causes the back-arc basin to close, resulting in deformation, metamorphism, and even melting of the strata deposited in the basin.

Mountain building

If the rate of subduction in an ocean basin exceeds the rate at which the crust is formed at oceanic ridges, a convergent margin forms as the ocean initially contracts. This process can lead to collision between the approaching continents, which eventually terminates subduction. Mountain building can occur in a number of ways at a convergent margin: mountains may rise as a consequence of the subduction process itself, by the accretion of small crustal fragments (which, along with linear island chains and oceanic ridges, are known as terranes), or by the collision of two large continents.

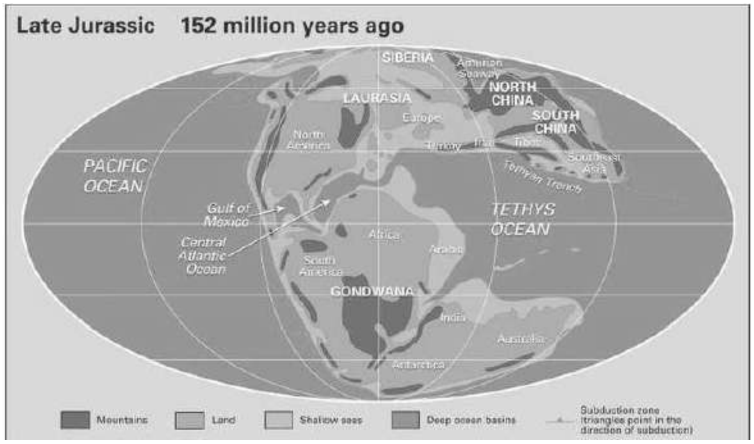

Fig. Jurassic paleogeography

Many mountain belts were developed by a combination of these processes. For example, the Cordilleran mountain belt of North America—which includes the Rocky Mountains as well as the Cascades, the Sierra Nevada, and other mountain ranges near the Pacific coast—developed by a combination of subduction and terrane accretion. As continental collisions are usually preceded by a long history of subduction and terrane accretion, many mountain belts record all three processes. Over the past 70 million years the subduction of the Neo-Tethys Sea, a wedge-shaped body of water that was located between Gondwana and Laurasia, led to the accretion of terranes along the margins of Laurasia, followed by continental collisions beginning about 30 million years ago between Africa and Europe and between India and Asia. These collisions culminated in the formation of the Alps and the Himalayas.

The distribution of landmasses, mountainous regions, shallow seas, and deep ocean basins during the late Jurassic Period. Included in the paleogeographic reconstruction are the locations of the interval's subduction zones.

Mountains by subduction

Mountain building by subduction is classically demonstrated in the Andes Mountains of South America. Subduction results in voluminous magmatism in the mantle and crust overlying the subduction zone, and, therefore, the rocks in this region are warm and weak. Although subduction is a long-term process, the uplift that results in mountains tends to occur in discrete episodes and may reflect intervals of stronger plate convergence that squeezes the thermally weakened crust upward. For example, rapid uplift of the Andes approximately 25 million years ago is evidenced by a reversal in the flow of the Amazon River from its ancestral path toward the Pacific Ocean to its modern path, which empties into the Atlantic Ocean.

In addition, models have indicated that the episodic opening and closing of back-arc basins have been the major factors in mountain-building processes, which have influenced the plate-tectonic evolution of the western Pacific for at least the past 500 million years.

Mountains by terrane accretion

As the ocean contracts by subduction, elevated regions within the ocean basin—terranes—are transported toward the subduction zone, where they are scraped off the descending plate and added—accreted—to the continental margin. Since the late Devonian and early Carboniferous periods, some 360 million years ago, subduction beneath the western margin of North America has resulted in several collisions with terranes. The piecemeal addition of these accreted terranes has added an average of 600 km (400 miles) in width along the western margin of the North American continent, and the collisions have resulted in important pulses of mountain building.

The broad, gentle pitch of the continental shelf gives way to the relatively steep continental slope. The more gradual transition to the abyssal plain is a sediment-filled region called the continental rise. The continental shelf, slope, and rise are collectively called the continental margin.

During these accretionary events, small sections of the oceanic crust may break away from the subducting slab as it descends. Instead of being subducted, these slices are thrust over the overriding plate and are said to be obducted. Where this occurs, rare slices of ocean crust, known as ophiolites, are preserved on land. They provide a valuable natural laboratory for studying the composition and character of the oceanic crust and the mechanisms of their emplacement and preservation on land. A classic example is the Coast Range ophiolite of California, which is one of the most extensive ophiolite terranes in North America.

These ophiolite deposits run from the Klamath Mountains in northern California southward to the Diablo Range in central California. This oceanic crust likely formed during the middle of the Jurassic Period, roughly 170 million years ago, in an extensional regime within either a back-arc or a forearc basin. In the late Mesozoic, it was accreted to the western North American continental margin.

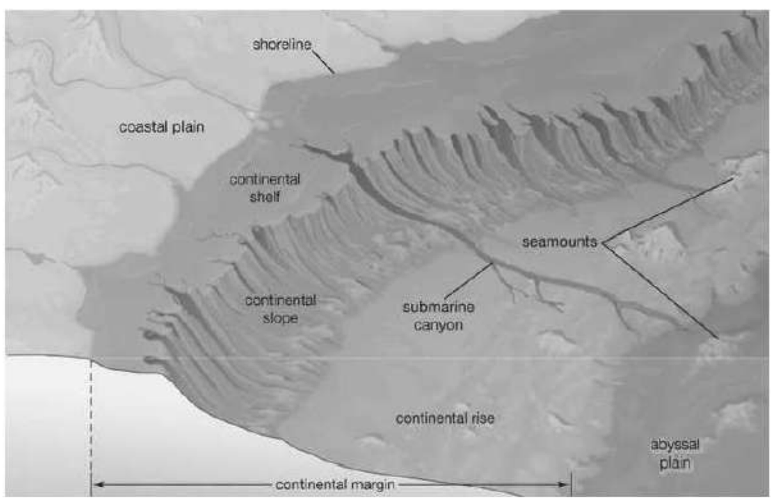

Fig. Continental margin

Because preservation of oceanic crust is rare, the recognition of ophiolite complexes is very important in tectonic analyses. Until the mid-1980s, ophiolites were thought to represent vestiges of the main oceanic tract, but geochemical analyses have clearly indicated that most ophiolites form near volcanic arcs, such as in back-arc basins characterized by subduction roll-back (the collapse of the subducting plate that causes the extension of the overlying plate). The recognition of ophiolite complexes is very important in tectonic analysis, because they provide insights into the generation of magmatism in oceanic domains, as well as their complex relationships with subduction processes.

Mountains by continental collision

Continental collision involves the forced convergence of two buoyant plate margins that results in neither continent being subducted to any appreciable extent. A complex sequence of events ensues that compels one continent to override the other. These processes result in crustal thickening and intense deformation that forces the crust skyward to form huge mountains with crustal roots that extend as deep as 80 km (about 50 miles) relative to Earth's surface, in accordance with the principles of isostasy.

The subducted slab still has a tendency to sink and may become detached and founder (submerge) into the mantle. The crustal root undergoes metamorphic reactions that result in a significant increase in density and may cause the root to also founder into the mantle. Both processes result in a significant injection of heat from the compensatory upwelling of asthenosphere, which is an important contribution to the rise of the mountains.

Continental collisions produce lofty landlocked mountain ranges such as the Himalayas. Much later, after these ranges have been largely leveled by erosion, it is possible that the original contact, or suture, may be exposed.

The balance between creation and destruction on a global scale is demonstrated by the expansion of the Atlantic Ocean by seafloor spreading over the past 200 million years, compensated by the contraction of the Pacific Ocean, and the consumption of an entire ocean between India and Asia (the Tethys Sea). The northward migration of India led to collision with Asia some 40 million years ago. Since that time India has advanced a further 2,000 km (1,250 miles) beneath Asia, pushing up the Himalayas and forming the Plateau of Tibet. Pinned against stable Siberia, China and Indochina were pushed sideways, resulting in strong seismic activity thousands of kilometres from the site of the continental collision.

Transform faults

Along the third type of plate boundary, two plates move laterally and pass each other along giant fractures in Earth's crust. Transform faults are so named because they are linked to other types of plate boundaries.

Fig. San Andreas Fault

The majority of transform faults link the offset segments of oceanic ridges. However, transform faults also occur between plate margins with continental crust—for example, the San Andreas Fault in California and the North Anatolian fault system in Turkey. These boundaries are conservative because plate interaction occurs without creating or destroying crust. Because the only motion along these faults is the sliding of plates past each other, the horizontal direction along the fault surface must parallel the direction of plate motion. The fault surfaces are rarely smooth, and pressure may build up when the plates on either side temporarily lock. This buildup of stress may be suddenly released in the form of an earthquake.

At the San Andreas Fault in California, the North American Plate and the Pacific Plate slide past each other along a giant fracture in Earth's crust.

Many transform faults in the Atlantic Ocean are the continuation of major faults in adjacent continents, which suggests that the orientation of these faults might be inherited from preexisting weaknesses in continental crust during the earliest stages of the development of oceanic crust. On the other hand, transform faults may themselves be reactivated, and recent geodynamic models suggest that they are favourable environments for the initiation of subduction zones.

Structure and Tectonics

Drainage may adjust passively to varying resistance of geologic materials, or it may be actively induced to follow a particular course by tectonism. Examples of the latter include faulting, as in the Ganges-Brahmaputra delta region. Growing folds and domes have affected drainage in the Colorado Plateau and central Australia. Subsidence has been important in the Mississippi and Pantanal regions.

Streams that emerge from mountain fronts onto surrounding plains display a fascinating array of structural and tectonic controls. Where mountain fronts are erosional because of a complex interplay of geomorphic variables, they may develop flanking surfaces of planation called pediments. Deposition at the mountain front produces alluvial fans because of the tremendous increase in width as a stream emerges from a mountain canyon. Examples include the Tian Shan, Kosi, and Pantanal areas. Passive adjustment to structure is a quality of nearly all the study areas.

Perhaps the most interesting situations, however, are drainage anomalies, where streams cut across structural zones. Some streams appear to take the most difficult routes possible through fold belts. In his studies of the Appalachians and the Zagros Mountains, Oberlander has applied the term "obstinate streams" to this phenomenon. The Finke River is an excellent example. The Colorado River provides other examples.

Channel Patterns

Rivers display a remarkable variety of channel patterns that are especially amenable to study using spaceborne remote sensing systems. The patterns relate to large-scale conditions of climate and tectonism that can only be appreciated on a global perspective. It is remarkable that, despite the geologic dominance of "big rivers", it is precisely those rivers that have received the least study. Experimental work by Schumm has done much to increase our understanding of channel patterns. Pattern adjustments, measured as sinuosity variation, are closely related to the type, size, and amount of sediment load.

They are also related to bank resistance and to the discharge characteristics of the stream. Many of the morphological dependencies of river patterns can be summarized in the following expressions:

$$Q_w \propto \frac{W, d, \lambda}{S}$$

$$Q_s \propto \frac{W, \lambda, S}{d, P}$$

These relationships are expressed by a large number of empirical equations treating the important independent variables, Q_w, a measure of mean annual water discharge, and Q_s, a measure of the type of sediment load. The dependent variables are the channel width, W, depth, d, the slope of the river channel, S, the sinuosity, P and the meander wavelength.

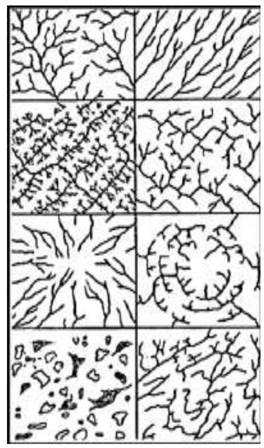

Fig. ***Major Types of Drainage Patterns***

The relationship of channel slope to sinuosity in an experimental river was elaborated by Schumm and Kahn. The data display a clear threshold phenomenon, in which steep low-sinuosity streams may change, somewhat abruptly, to somewhat less steep high-sinuosity streams.

The former comprise the bedload-type streams that yield braided patterns, whereas the latter yield the familiar meandering patterns associated with streams that transport a high suspended load. The shift between these two stable pattern configurations is illustrated by several study areas, including the Yukon, Kosi, Pantanal, Japurá, and Ucayali.

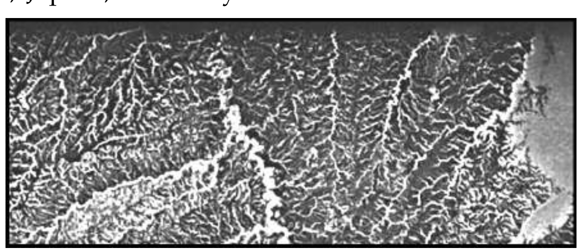

Fig. ***SIR-A Radar Image of Dendritic Drainage in Wast-central Columbia.***

Most of the image shows an area of dissected plains with a grassland cover that yields low radar return. The drainage pattern is strongly enhanced in radar return because the forested stream channels reflect the radar energy back to the receiver. On the basis of the foregoing experimental work, a variety of pattern classifications can be proposed. However, the immense complexity of natural fluvial systems appears to defy our present understanding.

Meandering Pattern

Meandering is the most common river pattern, and meandering rivers develop alternating bends with an irregular spacing along the valley trend. Such rivers tend to have relatively narrow, deep channels and stable banks. The system adjusts to varying discharge by vertical accretion on its floodplain and/or by lateral migration of its channel.

A vast complex of floodplain depositional features is associated with such rivers, as illustrated by the Mississippi River study area.

Braided Pattern

Braided rivers have channels divided into multiple thalwegs by alluvial islands. Braided rivers tend to have steeper gradients, more variable discharge, coarser sediment loads, and lower sinuosity than meandering streams. Their channels tend to be relatively wide and shallow. Braided patterns are "... developed depositionally within a channel in which the flow obstructions are sand and gravel deposited by the water moving around them". Midchannel bars are emplaced because of local flow incompetence.

The resulting braid channels formed by splitting the flow are more competent than the original channel for conveying the load downstream. Another way of describing braiding is that it is caused by channel widening that increases the boundary resistance of rivers with non-cohesive banks.

To maintain enough velocity for sediment transport in a wide, shallow cross section, the channel must divide and form relatively narrow and deep secondary channels through incision. Excellent examples of braiding occur in gravel-transporting rivers, such as Yukon, Colville, and upper Kosi. Braiding can also occur in sand-transporting rivers, like the Brahmaputra. The latter experience more frequent and more complex modification of original bar forms.

Anastomosed Pattern

Many multichannel rivers have relatively low gradients, deep and narrow channels, and stable banks. Such river systems have been termed "anastomosed". The terminology is a bit confused because "anastomosis" is a general designation for interconnected channelways whether in alluvial or in bedrock rivers.

Thus, Garner, following Bretz defined an anastomosing channel system as "... an erosionally developed network of channels in which the insular flow obstructions represent relict topographic highs and often consist of bedrock." Anastomosis is extensively developed in the Channeled Scabland.

Therefore, anastomosing patterns can be considered to be composed of multiple interconnecting channels separated by relatively stable areas of floodplain or bedrock. In contrast, braided patterns are single-channel, multiple-thalweg systems with bars of sediment or vegetated islands around which flow is diverted in the channel.

Excellent examples of anastomosed streams occur in the plainslands of east-central Australia. The Burke and Hamilton Rivers and the Cooper Creek study areas illustrate these arid-region varieties. Anastomosis also characterizes very large tropical rivers, such as those in the Amazon Basin. The Solimões and Japurá study areas illustrate such rivers.

Distributary Pattern

Distributary patterns occur where fluvial systems are spreading water and sediment across depositional basins. Two varieties are fans and deltas.

Fans develop in piedmont areas under the influence of both tectonic and climatic controls. Arid-region alluvial fans are constructed by infrequent depositional events that include both debris flows and water flows. Typical arid-region fans occur in the Tucson and Tian Shan study areas. Cold-climate alluvial fans occur in areas of glacial outwash and in periglacial regions. An excellent example is the Sheenjek Fan in the Yukon River study area. Humid-region alluvial fans are constructed by seasonal or perennial fluvial flows. The Kosi Fan of Nepal and India is an example from an area of active mountain building. The Pantanal study area illustrates some large fans in the savanna tropics of Brazil.

Deltas are the subject of another stage in this volume since most deltas involve the interaction of ponded water systems with sediment delivered to a river mouth. However, some basins of deposition in arid regions lack ponded water. Rivers entering these basins may produce typical deltaic morphologies, as in the case of the Niger River study area in Mali, West Africa.

Transitional Patterns

Five study areas from the Amazon Basin illustrate the complexity of tropical river systems. Many of these complexities arise because the fluvial system is not merely an entity that is totally adjusted to the vagaries of modern conditions. Rivers possess a heritage in which they inherit elements of ancient conditions.

Thus, old buried structures, relict alluvium, and progressive development contribute detail to the modern fluvial landscape. The understanding of modern rivers requires an understanding of their past history.

Tectonic Landforms

All continents are part of crustal plates and have two common components; cratons and folded linear mountain belts. Cratons are expansive, stable regions of low relief typically in the central part of the continent and made up of old igneous and metamorphic rocks buried under a relatively thin mantle of sedimentary rocks. The key part of the concept of a craton is stable. Although the old rocks forming the core may have been extensively metamorphosed in the geologic past, the craton has not undergone appreciable metamorphism or rapid uplift for several hundreds of million years. If other processes have not modified the sedimentary veneer of the cration, soils on the craton can be very old because of low relief and low erosion rates. The northern Great Plains of the U.S. and central Canada form the North American craton. Folded linear mountain belts are common on the margins of continents. Their occurrence is related to current or past collosions of plates. Rocks are typically extensively metamorphosed with intrusions of igneous rocks (Stone Mountain). Because relief is usually high, high rates of erosion prevent development of very mature soils. Both the Appalachian and Rocky Mountains are examples of folded linear mountain belts. The Appalachian Mountain chain formed about 300-350,000,000 years ago as the result of the collision of the North American and African plates.

The Piedmont of the eastern U.S. is part of this mountain chain, but erosion has worn down the mountains to the rolling landscape we see today. The Blue Ridge and other higher mountain chains further inland were folded and uplifted slightly later that the Piedmont and have not eroded as much. The Rocky Mountains formed from the collision of the North American and Pacific plates during the Tertiary epoch (65-1.8 million years before present (ybp)). Because these mountains are considerably younger than the Appalachians, they are higher and steeper. Other examples include the Himalayan Mountains, the Alps, and the Andes. Volcanism is also common during this type of mountain building. Volcanic deposits are common in the western U.S. and have been identified in the Appalachians and Piedmont.

Glacial Landforms

Climate change is the rule rather than the exception during the earth's history. Continental glaciers that covered much of the high latitude regions of the earth during ice ages have sculpted most of the landforms in these areas. During the Pleistocene epoch (1,800,000 to 12,000 ypb), there were multiple glacial advances and these glaciers covered much of the earth's higher latitudes.

In the U.S., glaciers extended as far south as St. Louis, MO during periods of maximum glaciation. Glaciers form a wide variety of landforms, and glacial deposits are widely variable. Till, the most widespread type of glacial deposit, was the material pushed, churned, and modified under the glacial ice.

An analogy is a huge bulldozer that pushes all the material across the landscape and the resulting deposit is a mixture of whatever was in the path. Till is typically poorly sorted with particle sizes that range from clay to boulders. Composition of the particles depends on what was present in the path of the glacier. Composition of till in the midwestern U.S. is different than that in the northeast because the rock types encountered by the glacier were different. One type of deposit not formed by the glacier but that is associated with glaciers is loess; silty eolian materials. Even during glaciation, there was winter and summer. During summer, glaciers melted and large quantities of meltwaters flowed away from the glacier in steams. Along with the water were large quantities of sediment. The result was large stream discharges with large quantities of sediment in major rivers such as the present day Ohio, Missouri, and Mississippi rivers. The river channels were very broad to accommodate the flow and large amounts of sediment were deposited in the channel and on the broad floodplain.

During winter, the glaciers melted less, and the flow and associated sediment load in the rivers was substantially reduced. The reduced stream flow left large areas of freshly deposited sediment uncovered by water and with no vegetation. This sediment was entrained by wind and deposited in the adjacent uplands where vegetation was present to trap the eolian sediments. Sands were deposited near the channel as eolian dunes. The finer silt and clay were transported further from the channel before being deposited as what we now refer to as loess. Loess deposits are extensive in regions that carried glacial meltwaters including the central U.S., northern Europe, and northern Asia. Because it is an eolian deposit, loess is a blanket that covers all existing landforms and rock types. Thickness is uniform on all surfaces locally, but thickness of the loess deposit decreases with distance away from the river source. Near the river valley, the loess may be many meters thick and gradually thins with distance from the source. Because the Midwest U.S. had numerous stream sources, most of this part of the U.S. is has a loess cap that is <1 to several meters thick and blankets whatever rock or sediment that was present (till, sedimentary rocks and deposits).

Along the Mississippi River, whose ancestor drained most of the glaciated region of the U.S., thick loess deposits occur as far south as central Mississippi. The loess is several meters thick near the river and thins away from the river to the point that it cannot be identified more than 80-100 km from the river. Soils associated with glacial landforms have a known age which cannot be older than the time of the glacial advance. The last glacial maximum in North America was about 12,000 ybp. Thus, soils in the central U.S. formed in either glacial deposits or loess are younger than 12,000 years. Soils formed in loess are silty and typically have <10 per cent sand.

"Marine" Landforms and Deposits

The term "marine" deposits is somewhat of a misnomer because most of the deposits we commonly refer to as marine as actually deposited at the sea-land interface. Thus, coastal deposits or continental shelf deposits would be a better term. Most continents have coastal regions, typically with relatively low relief, with unconsolidated sedimentary materials that were deposited in a variety of coastal environments including beaches, dunes, marshes, and deltas similar to those landforms that are currently on the coast.

As stated earlier, climate change is the norm not the exception. Commonly associated with climate change is a change in sea level. Using Georgia as an example, during the Tertiary epoch, the ocean covered at least the southern half of the State, and some evidence suggests the sea level was as far north as Gainesville at one time during this period (Athens was under water). During the Pleistocene epoch, sea level was as much as 100 m lower than it is currently. In coastal plain regions (southern ½ of Georgia), particle size and composition of sediments (soil parent materials) vary with the environment in which they were deposited. For example, soils developed on ancient beach dunes are very sandy. Soils developed in ancient marshes are clayey and the list goes on. Since the sediments were transported to the coast by rivers, mineral composition of the sediments is strongly influenced by the mineralogy of the soils and sediments in the watersheds of the streams.

Limestone is another type of "marine" deposit that is composed of calcite ($CaCO_3$) or dolomite ($CaMg(CO_3)_2$). Most limestone was precipitated by marine organisms on the continental shelf some distance from the shoreline (think of coral reefs). Thus, limestone often contains a minimal amount of silicate minerals (quartz, feldspar, clay minerals, etc.). During the Cretaceous epoch (age of the dinosaurs; 146-65 million ybp), warm shallow seas (prime environment for carbonate formation and limestone deposition) covered extensive areas of North America (much of the central part of the continent) and thick limestone deposits occur in these areas. As limestone weathers and soil begins to form, the calcite and dolomite dissolve and are leached from the soil in ionic form. Thus, limestone derived soils are formed from the non-carbonate residues which are often clayey. Thus, soils over limestone in humid regions are often clayey and are also often red. In arid regions, limestone derived soils often have accumulations of calcium carbonate in the subsoil because of incomplete leaching of the carbonate. A question to ponder; if a soil derived from limestone that contained 5 per cent silicate residue is 2 m thick, how many feet of limestone were dissolved to form the soil?

Plate Tectonic Forces

Plate tectonic is the idea that plates carry the continents and are great slabs of solid material that make up the ocean floor. Plate tectonics comes from the Greek word,"tektonikos" meaning "builder." It has been determined that there are about 20 rigid plates that are in slow, continuous motion. Some continents move at a rate of 1/2 to 4 inches per year which is directed by heat driven convection cells in the molten rock deep below the crust. As they move, they carry the continents and ocean floor.

In the late 1800's, Alfred Wegener, a German physical geographer, used spatial analysis to propose the continental drift hypothesis. Wegener studied the outlines of the continents and suggested that the existing land masses had been united at one point in the earths early history. He called his theory"die Verschiebung der Continent" meaning "continental displacement." His idea stated that these stable, immovable continents were mobile with the help of the tectonic plates. With further research his theory was accepted, but not until 60 to 70 years later. The Earth is made up of three layers: the crust, the mantle and the core.

The crust is a thin (15 mile) layer covering the outside part of the earth. The second layer is the mantle which is 1,800 miles thick. The crust and the upper mantle make up what is called the lithosphere. This lithosphere is 60-90 miles below the continents and 40-50 miles below the oceans. The plates in the plate tectonic theory are the lithosphere. The continental crust is less dense or lighter than the oceanic crust and "floats" above it. The base of the lithosphere is called the asthenosphere. Here is where the lithosphere is unattached from the mantle and moves around, mostly by gravity and thermal differences in the mantle.

The core, the third layer, is 1,000 miles thick. The core and the mantle are made of hot molten rocks, but the core is much hotter than the mantle. Below the 15 mile crust there is an increased amount of heat. It is believed that this heat is 'left over' from the formation of the earth and decaying radioactive material is fueling the fire. In fact, it is a possibility we could use the energy from this heat to fuel our lives if we were to run out of oil. How does this heat cause the plates to move? The earth's crust is cold, the mantle is hot and the core is even hotter, thus providing us with the explanation. In order to equalize these temperatures, convection cells are formed. Two types of rotation are produced by these convection cells.

Propelled by these heat convection cells, these plates move very slowly; one to four inches per year. (A couple of inches per year isn't much, since the earth's history is measured in millions of years). 225 million years ago, it is postulated that a giant continent called Pangaea ('all-earth') existed.

This giant continent remained until about 135 million years ago, when it began to break up during the Mesozoic time. The break up consisted of India detaching itself from Africa and Antarctica and headed into the Indian Ocean. A giant mountain range is formed where the Australian-Indian Plate is pushing into the heart of Asia.

Additional evidence supporting the continental drift theory is supplied by the amphibian and reptile fossils that are spread out among the widely separated continents. And the evidence of polar wandering and evidence of magnetic field reversals locked into oceanic basalt samples. There are only three ways plates can interact while moving. This, in turn, causes there to be only three types of boundaries which are produced by different stress fields. The first boundary is called divergent. This is a tension or a stress that pulls the plates apart.

Divergent boundaries cause mid oceanic ridges. Some other common characteristics are high heat flow, mild volcanic activity, and shallow earthquakes. The second boundary is called convergent and this is a compression, or a stress that can shorten or compresses the plates. Convergent boundaries cause mountain ranges to develop. The Himalayas for example, were formed when the plate carrying India collided with the plate carrying Eurasia. This continental collision is still active and moving at a rate of 3 to 4 inches per year as the India plate is pushed under the Asian plate and the mountain continue to grow. Strong earthquake activity is very common with areas of convergent boundaries as well.

The third type of boundary is a transform boundary and this is when the plates slide past each other along faults causing mid-oceanic ridges and trenches. One plate may be forced down into the mantle under the other plate. When this occurs, a deep oceans trench forms.

The largest ocean form is the Mariana Trench in the Pacific Ocean southwest of Guam. Another example is the San Andreas Fault with is between the North American plate and the Pacific plate. Here, earthquakes are common but not volcanoes, and the earthquakes tend to outline the major plates.

The earth is constantly being shaped by the dynamics of the tectonicactivity and plates motion. Some of the present day's biggest mountains and ocean trenches are examples of the great power of these plates in motion.

'Hot Spots' is a generally accepted term used to explain the formation of islands in the middle of the Pacific such as the Hawaiian Islands, the Line Islands, and the Tuamotus. The Galapagos Islands are similar in fashion, though not as aligned, but are located off the coast of Ecuador. A Hot Spot is caused by the magma that rises or plumes from the core to the surface causing volcanoes by penetrating the mantle.

As the plate moves, it carries along the volcano that was formed. In it's place, a new one begins to form from the sea floor, while the hot spot stays in one place. Islands form in a "chain" as a result.

The Hawaiian Islands get younger from east to west in the chain. On the island of Hawaii, which is still over the hot spot, volcanoes remain very active. Other events related to this activity include earthquakes, volcanoes, and geothermic activity. Earthquakes are caused by abrupt easing of strains that have been built up along geologic faults and by volcanic action. The result of this is movement in the earth's surface. These vibrations can be felt in the locally affected areas and measured by scientific devices around the world. These plate cycles may also form volcanic activity.

For example, the Pacific Ocean is surrounded by a nearly continuous plate-collision zone, called the 'Ring of Fire', here Volcanoes are the results of the instability of this zone. Japan lies near the colliding edges of three plates, hence, earthquakes and volcanoes are a constant threat to the islands population.

Tectonics, Structures and FT Thermochronology

Structural geology is the study in theory, in the laboratory and in the field of mineral and rock deformation at scales ranging from intracrystalline to continental. Tectonics is the study of the construction of the Earth - how large-scale processes of rock formation and deformation interact to create the planet. Research at Penn focuses on regional aspects of these sciences. Emphasis is placed on field study and laboratory analysis to develop regional deformational histories to aid in understanding the assembly of geologic regions. Investigations examine settings as varied as modem and ancient convergent plate boundaries (the Mariana forearc and the California Coast Ranges) and intracontinental deformation belts (basement uplifts in Iberia and in the Rocky Mountains). Within these regions, inquiries have included experimental deformation of serpentinite muds, studies of basalt geochemistry, fission-track analysis of basement and cover rocks, and seismic and gravity interpretation. Extensive collaboration with colleagues from other leading research institutions provide our students with access to a wide range of expertise, equipment and perspectives. Among others, these institutions include the Lamont-Doherty Earth Observatory, the Smithsonian Institution, the University of Hawaii and the University of Lisbon.

FT Thermochronology: Fission-track thermochronology is one of the newest and most powerful tools geologists use to reconstruct tectonic and thermal histories of diverse geologic terranes. The method has been applied to many fields of geology, thus enhancing positive interactions between scientists in different disciplines of earth and planetary sciences. Research in this program involves the deciphering of tectonic and thermal histories of rift margins and basins, orogenic belts and sedimentary basins, and dating of meteorite impact events

Structural Geology and Tectonics:

1. Intracontinental Deformation: Spain, Portugal and the midcontinental U.S. Comparative investigation of the intracontinental consequences of collisional orogeny at

continental margins. Such collisions produce deformation far into the continent when basement blocks are shifted along old faults, commonly those related to failed rifts.

2. Late Mesozoic Paleogeography of the U.S. Pacific Coast: California & Oregon. Current models do not explain the presence of a previously unmapped ophiolitic milange unit lying within the forearc of the Franciscan subduction complex but outside the complex itself. The modern Mariana subduction zone appears be a good actualistic analogue for many aspects of the fossil Late Mesozoic one along the U.S. Pacific coast. Investigations at Penn focus on testing this comparative model, and currently include several field projects along a spectacular section of the coast in southwestern Oregon.
3. Tectonics of the California Coast Ranges: Compressional stresses between the Pacific and North American plates that are not resolved by transcurrent motion along faults of the San Andreas system are being taken up by active thrust faulting and related folding, forming a variant of Valley-and-Ridge geometry containing exotic rock types such as ophiolites and forearc-basin subsea-fan rocks. Work in this area includes fission-track studies (in collaboration with Dr. Omar) and studies of fault geometry and timing.

Dr. Omar uses FT and thermal modeling to reconstruct tectonic and thermal histories of the following geologic terranes:

1. The Red Sea Rift: Timing, geometry, and extent of rift flank uplift and its relationship to subsidence history of the rift basin. Ultimate objective is to obtain a greater understanding of the interaction of tectonics, geologic structure and erosion at the Red Sea rift margin.
2. Atlantic Margin Rift Basins: Newark and Taylorsville basins. Evaluation of the timing, spatial distribution, and migration of crustal-scale fluid-flow, a poorly understood phenomenon, the effects of which are only now becoming apparent.

3. Southeast Brazil: The goal of this project is to allow definitive limits to be placed on the timing of uplift and denudation of the southeast Brazilian topography and mechanisms for their formation.
4. Rocky Mountains: the main objective of this project is to determine the tectono-thermal history of individual basement blocks and intervening sedimentary basins in order to decipher the tectonic and deformation history of the western USA.

7

Quantitative Morphology

GEOMORPHOLOGICAL MAPPING AND GEOMORPHOMETRY

Most of the interest in geomorphological mapping has centered on the development of various mapping systems for use in environmental management. The most detailed systems have been developed in Europe, where different countries utilize different procedures. Despite attempts at international standardization, the major problem remains the correlation of various mapping schemes.

Global remote sensing has great promise for generating geomorphic maps at large spatial scales. The use of the SIR-A system to define terrain categories for geomorphic mapping in Indonesia. The terrain types are recognized through interpretation of radar interaction with the ground surface, especially the surface roughness, vegetation, and topography. The radar information is made even more remarkable because the mapped areas are characterized by dense tropical forest cover and persistent clouds.

Future Directions

In 1982 in Vienna, the United States proposed at the United Nations Conference on Peaceful Uses of Outer Space that an international cooperative research programme be organized to understand the Earth as a system. This programme was initially named "Global Habitability" and was formulated to involve the central role by NASA in the observation of system parameters and changes.

In a sense, the programme constitutes a "mission to planet Earth" in which spaceborne remote sensing is applied to studies of dynamic processes in the atmosphere, biosphere, geosphere, and hydrosphere. The Global Habitability concept will probably be integrated into the broader efforts of the proposed International Geosphere-Biosphere Programme to be coordinated by the International Council of Scientific Unions.

In his keynote address to the World Conference on Earthquake Engineering, the President of the National Academy of Sciences proposed an "International Decade of Hazard Reduction". A major component of such an initiative must be the global analysis of hazardous geomorphic processes. What do such initiatives mean for geomorphology?

Methods and Principles of Geologic Mapping

This brief account of the early geologic maps gives us glimpses and hints of the methods and principles used in geologic mapping. We should now summarize these principles, consider their validity, and show how they are applied in actual field work.

On William Smith's geologic map of England, lines representing the contacts between different rock formations are drawn for distances equivalent to hundreds of miles on the ground. Yet, in tracing an individual contact for 100 miles, Smith probably found, on the average, less than 50 exposures where the actual contact could be seen on a clean rock face. How, then, can his map record the real distribution of the rocks? Can it be anything but a guess?

The eyes of a geologist are no more capable of seeing the bedrock through a cover of soil than those of any other observer. How, then, can the geologist make inferences about the position of the bedrock underground that will withstand objective tests, such as those provided by digging wells and sinking mine shafts?

The succession must be pieced together from scattered outcrops. Although in an area of a few square miles a geologist may find only one or two outcrops in which he can see the actual contact between two formations, he will doubtless find 100 or more outcrops composed entirely of rock belonging to one or the other formation.

He thus has some clue as to the position of the contact with respect to each outcrop. The problem is roughly analogous to drawing a contour line to conform to the elevations determined at 100 or more control points. To use scattered rock outcrops, however, requires that the individual formations be correctly identified (correlated) from outcrop to outcrop, and this is not always easy.

Four Fundamental Postulates of Geologic Mapping

There are four fundamental postulates that underlie the making of a geologic map. The third is merely a commonsense deduction from these two, and, like them, was first stated by Steno: A water-laid stratum, at the time it was formed, must continue laterally in all directions until it thins out as a result of non-deposition, or until it abuts against the edge of the original basin of deposition. This is the Law of Original Continuity. An important corollary of this law, not fully appreciated by Steno in 1669, but well known to the geologists of France and England at the beginning of the Nineteenth Century, may be stated: A stratum which ends abruptly at some point other than the edge of the basin in which it was deposited, must have had its original continuation removed by erosion, or else displaced by a fracture in the earth's crust. These four postulates—

- Superposition (the higher bed is the younger);
- Original horizontality (stratification planes are formed roughly parallel to the earth's surface);
- Original continuity;
- Truncation by erosion or dislocation—are the basis for many of our interpretations of the relative relations of strata.

They are not absolute rules that can be rigidly applied. For instance, some beds, once horizontal, have been highly tilted and even overturned by movements of the earth's crust , so that a stratum formerly beneath another may now lie inverted on it; other strata, as at the front of a steep delta, may have been deposited on appreciable slopes; the edges of many landslides end abruptly instead of thinning to a feather edge.

Such exceptions, however, are not common and can generally be recognized easily by the geologist. How do we correlate between scattered outcrops when we are in the field making a geologic map?

Correlation of Rock Outcrops

In a ravine on a grassy hillside we may see a bed of clay with well-marked horizontal stratification; we assume it continues horizontally into the hill at the same elevation, for how else can its horizontal stratification be extended?

If we go 200 yards along the same contour, without seeing an outcrop, and then find a clay bed at the same elevation in another ravine, we may suspect it is the same bed. If both are gray, the probability is heightened. If both show lines of nodular lumps (concretions) along the stratification planes, and if the size and spacing of these lumps is about the same, we are still more confident of our correlation.

If both rest on red limestone, both are overlain by fine-grained brown sandstone, and both contain the same kinds of fossils, we become practically certain of the correlation. We are now justified in assuming that the bed is continuous beneath the soil between the two exposures.

If we go on a little farther and come to a deep ravine in the hill which shows a very large outcrop containing not a few feet of strata but several hundred, and if in this section there is only one clay bed with features identical to those we saw in the two small outcrops, we have still further evidence that our mapping is correct. We can now map the clay, for the contacts of the bed with those above and below will parallel the contours, as they are horizontal planes along the top and bottom of the clay bed.

We have, furthermore, gained information on the thickness and relations of the red limestone below and the brown sandstone above the clay bed, because in this larger outcrop they are exposed through a much greater thickness. For example, we may find that the brown sandstone is 180 feet thick and is overlain by a thick bed of distinctive black limestone crowded with fossils.

If, now, we go on to still another outcrop and find here the clay bed is a little thinner than in the last outcrop, if instead of resting on red limestone it now rests on a pink limy sandstone, and if it has a pebbly green sandstone instead of the fine-grained brown sandstone above it, we would be less certain of our correlation, though we would not regard it as impossible.

To check on the correlation, we might go to a lower elevation and study the relations of the sandstone in various outcrops. If we found it gradually varying from limestone to sandy limestone and then to pink sandstone, our correlation would be strengthened.

It would be further confirmed if we walked up the hill and found that the pebbly green sandstone above the clay is overlain by a black limestone full of fossils exactly like those in the limestone that overlay the brown sandstone in the large exposure.

Thus in geologic mapping there is invariably an element of judgment. Some correlations are certain, others are reasonably sure, and for still others there may be a reasonable doubt. As to the doubtful ones, two geologists may disagree, just as two equally qualified physicians sometimes disagree in the diagnosis of identical, but not sufficiently definitive, pathological symptoms.

But nearly all such differences in interpretation have to do with minor features in a stratigraphic succession. Thicker groups of beds commonly exhibit enough peculiarities to lead any two careful observers to identical conclusions.

Geologic Sections

Geologic sections are commonly used along with geologic maps in all economic applications of geology. A geologic section represents the way the rocks would appear on the side of a trench cut vertically into the land surface. It resembles a geologic map in representing the projection of the rock boundaries to a surface, but this surface in a cross section is vertical and not the land surface of the earth.

As an example, let us make a geologic section of the horizontal clay bed just mentioned, taking our data from a geologic map showing the relations of the clay in a deep branching ravine where the beds are well exposed.

The completed section appears below the geologic map, together with the construction lines used in drawing it. The geologic section portrays the strata as they would appear on the wall of a trench dug along the line A-B of the geologic map.

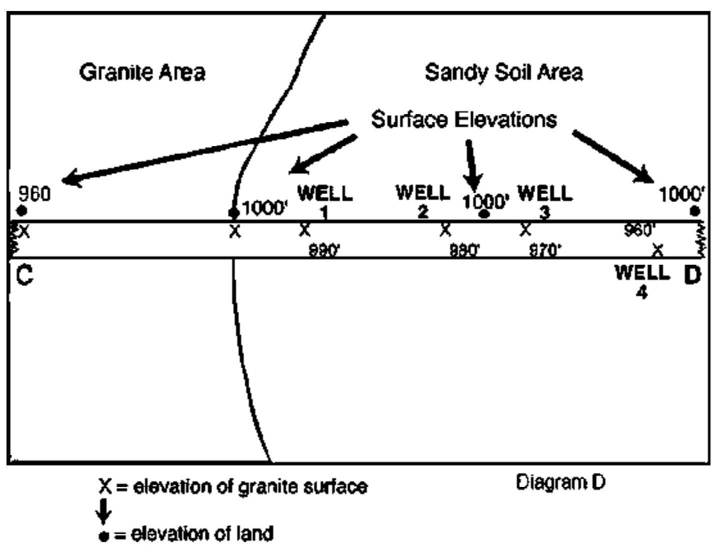

Fig. ***Construction of a Geologic Section from a Geologic Map.***

The section was readily drawn by making use of the contour lines on the map and relating the outcrops to them. First, it was decided to make both the horizontal and vertical scales of the section the same as that of the map, in order to avoid distortion..

Then perpendicular lines were dropped from A and B (by means of a right-angle triangle) to the place below the map chosen for drawing the section. Then a vertical scale showing the range in elevation from 1120 to 1,260 feet above sea level was laid off on the line projected from B.

Horizontal lines were drawn across the section, from the point representing the elevation of the respective contours from 1,140 to 1,260. Then a perpendicular was dropped from each point of intersection of line A-13 with a contour line on the map, to the horizontal line on the section representing the appropriate contours. Thus, points were plotted on the section to guide the sketching of the *surface profile*, the irregular line A2 -B2 of the section.

Now perpendiculars could be dropped to the profile from the intersection of the map line A-B with each contact of the clay stratum. There are five such intersections shown along the line of section—three on the top of the clay stratum, two on its base.

Horizontal lines drawn through the projections of these points on the surface profile complete the cross section and show the relations of the clay bed to the underlying limestone and overlying sandstone.

Many practical uses can be made of geologic sections. If the clay in the bed is suitable for brickmaking, we can determine from the section how much overburden must be removed at any point along the section line so as to get down to the clay. If water is found in wells just above the clay, the section tells us how deeply we must drill at any point along the section in order to strike it. Mine shafts, tunnels, and drill holes all test the validity of our maps and sections and of the postulates underlying their construction.

Geomorphometry

Geomorphometry is the science of quantitative land surface analysis. It gathers various mathematical, statistical and image processing techniques that can be used to quantify morphological, hydrological, ecological and other aspects of a land surface. Common synonyms for geomorphometry are geomorphological analysis, terrain morphometry or terrain analysis and land surface analysis. In simple terms, geomorphometry aims at extracting (land) surface parameters (morphometric, hydrological, climatic etc.) and objects (watersheds, stream networks, landforms etc.) using input digital land surface model (also known as digital elevation model) and parameterization software. Extracted surface parameters and objects can then be used, for example, to improve mapping and modelling of soils, vegetation, land use, geomorphological and geological features and similar.

Although geomorphometry started with ideas of Brisson (1808) and Gauss (1827), the field did not evolve much until the construction of the first DEM. With the rapid increase of sources of DEMs today (and especially due to the Shuttle Radar Topography Mission and LIDAR-based projects), extraction of land surface parameters is becoming more and more attractive to numerous fields ranging from precision agriculture, soil-landscape modelling, climatic and hydrological applications to urban planning, education and space research. The topography of almost all Earth has been today sampled or scanned, so that DEMs are available at resolutions of 100 m or better at global scale. Land surface parameters are today successfully used for both stochastic and process-based modelling, the only remaining issue being the level of detail and vertical accuracy of the DEM.

In hydrology, behavioural modelling is a modelling approach that focuses on the modelling of the behaviour of hydrological systems.

The behavioural modelling approach makes the main assumption that every system, given its environment, has a most probable behaviour. This most probable behaviour can be either determined directly based on observable system characteristics and expert knowledge or, the most frequent case, has to be inferred from the available information and a likelihood function that encodes the probability of some assumed behaviours. This modelling approach has been proposed recently by Sivapalan et al. (2006) in watershed hydrology.

DEM AND DIGITAL GEOMORPHOMETRY

Points and DEMs

To facilitate interpretation, the 2D digital images acquired with the integrated camera can be used to give a true colour value to the lidar points; this is especially useful as a low-resolution preview of the entire area. Using the image calibration, position and orientation, perspective projection is used to project each 3D laser point into the corresponding registered image. The colour of the image pixel is queried and assigned to the projected 3D point.

Registered point clouds may be combined and manipulated in any way desired by the user to gain the geological interpretation or geometric measurements required by the application. Points of interest may be selected from the vast overall point cloud, such as only the points corresponding to a single layer, by drawing a polygon round the desired points to form a new 'mini' point cloud. Quantitative measurements may be carried out, from simple distances between two points, fitting best-fit planes through the point clouds, and calculating strike and dip measurements from selected points. Available software dictates the features that are available.

Generally, two forms of the lidar data are of most use in outcrop studies: the point cloud data and DEMs (also called surfaces or meshes) formed from the original points. Use of the point cloud provides the highest accuracy, as all the collected data are available to interrogate. However, the sheer quantity of 3D points is often a barrier to working with the raw data, as the features of interest may be massively oversampled with respect to the topography of the outcrop. An example is an outcrop face where the geometry of bed boundaries is to be digitised. Although the vertical topography (cliff face) may be relatively smooth, the laser scanner does not account for this and records data equally on rugged and smooth surfaces, resulting in a large quantity of points. To record the bed boundaries, it is the line geometry that is most important, rather than the surface roughness, which may be more a product of weathering or manmade quarrying. To save on the computer resources required to load large point clouds (a merged dataset may contain more points than can be loaded), decimation of the raw point data and creation of a DEM is valid. Texturing, or draping, the digital imagery from the camera onto the created DEM results in a photorealistic model that is of extreme value for interpretation, visualisation and education. Use of a textured virtual outcrop model also addresses another problem with using points alone: that of accurately identifying fine-scale features. Although the point cloud may have very high resolution, at a certain interpretation level, on zooming in, the point spacing becomes too large to accurately identify fine-scale features. A textured model uses the higher resolution and continuity of the digital imagery to 'fill' the gaps between the points, making interpretation easier.

DEM Creation from Point Clouds

Creating a virtual outcrop model involves finding a best-fit surface through the raw points, to produce a triangulated mesh. This is similar to methods used for aerial terrain data, where a point set is triangulated to form a grid or triangular irregular network (TIN) model, but different from the 'surface gridding' approach utilised by most (but not all) seismic data and subsurface reservoir modelling packages commonly used today. The advantages of using TINs have been discussed in detail by McCullagh (1998). The major difference, and indeed difficulty, with the data acquired from a terrestrial laser scanner is that the point clouds are truly 3D, having points on vertical and overhanging surfaces. This is in stark contrast to conventional aerial data, where it is usual to have only one point for each (x, y) DEM point. It is easy to form a DEM from such a point distribution, as a 2D Delaunay triangulation finds the best criteria for triangle creation automatically (e.g. McCullagh 1998).

However, running such an algorithm on 3D lidar points will result in an erroneous surface model, where points of all elevations are connected on 2D adjacency, irrespective of range or obstructions. Three-dimensional surface reconstruction is therefore a complicated procedure, with much research in the fields of computing, and it is yet to be solved to allow fully automated algorithms. Nevertheless, software is available that can make DEMs from input scans. The process is facilitated by using the geometry of the collected data as an aid in determining the mesh; triangulation of a single scan is a simple matter, because the point cloud is a rough grid of points with angles and distances known relative to the instrument position. Therefore a 2D triangulation is possible, with a maximum edge length filter to prevent connecting points that are far apart in range. Integrating several scans is more problematic, as it is difficult to automatically triangulate complex overlapping point clouds, and editing is usually required.

Some preparation of the point cloud data is usually necessary prior to mesh creation. Because the original dataset is very dense, and may contain vegetation and other non-geology points, a pre-processing stage is carried out. Points may be manually selected and removed from each of the point clouds, a potentially time-consuming procedure if many unwanted points exist. Some automation is possible. Filtering of the point cloud may be carried out according to the intensity of the returns, where enough resolution of the digitised signal is provided by the instrument to distinguish different materials. Generally, the more noise caused by vegetation or registration errors there is in the point cloud, the worse the success rate of automatic meshing procedures will be. To combat this, preparation of the point cloud for meshing may be continued by smoothing and decimating the overall point set; because of overlap between different scans required to obtain full coverage of the outcrop, much redundancy may be present, resulting in an uneven point density. The point cloud may be decimated on surface curvature, or by using an octree, where a 3D grid is formed and each cell is populated using an average value of all points found within the cell volume (e.g. Girardeau-Montaut et al. 2005). The result is a more regular point cloud, with little loss of accuracy. In the experiences gained so far, it is possible to significantly reduce the original point cloud, whilst maintaining an accurate mesh. Indeed, aggressive decimation must be carried out to make models that can be loaded and visualised at interactive frame rates. This is a decision-making process that must be evaluated for each model, and for the size of the area and geological features that are being studied. This is a limitation with current hardware that will improve with time, but in general it can be said that if the utmost accuracy is required, the point clouds must be broken down into smaller areas to be able to be meshed and textured.

Some mesh editing is usually required after automated processing has been used, to fill holes and remove errors where the algorithm was unsuccessful. Although only a subset of the original scan data may have been used to make the mesh, this may still represent a significant number of points, leading to a large number of triangles (around double the number of input points). Loading textured outcrop models is intensive on PC resources, and the RAM and graphics card specification becomes critical. Therefore, a final mesh editing stage is intelligent decimation to reduce the number of triangles in areas of low surface roughness by representing them with larger triangles, whilst keeping smaller triangles in detailed areas. A paradox of outcrop modelling is that often the smoothest surfaces are of most interest (i.e. a vertical outcrop face), so care should be taken during decimation that these areas are not over-simplified.

Virtual Outcrop Formation

The final stage in lidar data preparation is the integration of the triangle meshes with the registered digital camera imagery to form a virtual outcrop model, and the quality control of that model. The images are used as textures, which are mapped (or 'draped') onto each triangle in the mesh, based on the projection of each triangle vertex into the digital image. Because a large number of images may be available, criteria for choosing the most suitable image part for each triangle must be used. These criteria are based on the direction the camera is pointing with respect to the triangle orientation, distance from the camera to the triangle, and quality of the image texture itself. Different lighting conditions or poor image exposures will severely reduce the visual impact of a virtual outcrop model, and may make interpretation difficult. A key issue is when adjacent triangles are textured with image data taken from very different angles or distances and under different lighting conditions. This highlights the triangles and distracts from the overall virtual outcrop quality. Therefore, attention to image acquisition while in the field is extremely important, more important than scan acquisition, which is light-independent, and should be given priority to achieve a suitable end product without performing additional and time-consuming manual image enhancement. Once suitable images have been chosen and optionally enhanced, the triangle texture mapping procedure is largely automatic, resulting in a 3D photorealistic model.

Interpretation and Measurement

The processed virtual outcrop model (or point clouds) can be visualised in a 3D environment, where the user is able to change the virtual camera viewpoint to suit their needs. In this way, inaccessible areas in the field become accessible, and interpretation and measurement can be performed, which can then be integrated with traditional field data. Although visualisation is extremely valuable, it is the potential for adding quantitative information that makes spatial data collection technologies so useful. The aim for many studies is to model an outcrop analogue within subsurface, geocellular reservoir modelling software, requiring accurate geometry from the lidar data. It is relatively simple to interpret the geology directly onto a virtual outcrop model, defining interest points or linear features as required. These features can be later imported to the reservoir modelling software, to be used in the generation of surfaces that provide the framework for the geocellular models. For example, a surface may be tracked in three dimensions across a wide area (and multiple virtual models), using the GPS absolute positioning to relate the different sections in space. The surface is represented by 3D lines, and the geology reconstructed by interpolating and extrapolating the line points to fill in the geometry between exposed areas. The lidar geometry therefore provides a basis for building geological models that allows more detailed studies to be carried out, on smaller-scale features.

The aim was to map the clinoforms, to extract numerical data on their thickness and lateral extent, and to use the mapped data as a framework for building 3D geocellular models. The geocellular models were built in software typically used for modelling subsurface hydrocarbon reservoirs, which allow fluid flow to be simulated under reservoir conditions. Thus the detailed geometries observed in the outcrop can be used as an analogue for subsurface cases where geometric data are sparse or lacking. The sedimentology and stratigraphy of the Perron Sandstone have been extensively documented by Ryer (1981) and Anderson & Ryer (2004).

The study area was 2 km^2 and includes around 3 km of near-continuous cliff section located at different orientations, ensuring both close to depositional strike and dip coverage. The data for the virtual outcrop comprised 18 scans collected over 5 days. The total dataset included c. 63 million raw points, 470 registered images, and an additional 45 images taken separately from the scanner. The point data were meshed and textured to form a virtual outcrop model that was then used for interpretation. Clinoform contacts were digitised on the virtual outcrop, with the user able to rotate and translate the model until the best orientation was found for seeing the geology. The result of this was a set of detailed line features (vertical separations from decimetre to metre scale), which were analysed and imported to geocellular modelling software for building surfaces using the lines as constraints. These surfaces, combined with conventional sedimentary logging carried out in the field, were used to model volumes representing the true clinoform geometry, in greater detail than previously possible. Such a study demonstrates the contribution of lidar techniques, as the vertical outcrop face would have made in situ measurements hazardous and extremely time consuming. Here, the lidar acquisition could be combined with logging and interpretation to build up a truly detailed dataset in only a few field days. The curved nature of the outcrop and the accurate 3D mapping with lidar allowed the clinoform surfaces to be accurately recreated in a way that is not possible in two dimensions. The dataset could be explored in much greater detail back in the office and, if necessary, a further field campaign could be carried out at a later date to verify interpretations made on the virtual outcrop.

Building surfaces from interpreted line features is only one measure that can be captured using the virtual outcrop model. Additional examples are strike and dip measurements, plotting fracture networks and obtaining cross-sections of the outcrop surface. Again, the application defines the required products.

Software for Lidar Processing and Interpretation

The above discussion has focused on the generalities of lidar processing, without being specific to the software tools that are available. Such software tools are transitory, and description of the various merits of each is out of the scope of this chapter. A number of commercially available software tools exist for processing lidar data, although it should be noted that these packages have not been developed specifically for geology applications. Indeed, most have not been developed for terrestrial survey applications and the long-range laser scanners that are of most use in geology, and instead have been driven by industrial applications (such as scanning manufacturing components in the case of PolyWorks by InnovMetric, one leading package). Such close-range applications are normally devoid of much of the data noise found in real-world environments, and therefore some algorithms and procedures are not yet perfected, making the development of a processing workflow not entirely as simple as selecting an off-the-shelf package. Software usually offers import of native scanner formats, registration by targets and surface matching, mesh creation (with variable success) and editing, as well as texturing using imagery. Additionally, some software (PolyWorks by InnovMetric; RiSCAN Pro by Riegl) allows digitisation of computer-aided design (CAD) features, which can be used for interpretation and input to modelling.

Because of the large amount of data that are generated, more sophisticated handling and visualisation strategies are required, provided by the dedicated software packages. The true 3D nature of point clouds and virtual outcrop models means that this form of data is not suitable for inputting to most geographical information systems (which are inherently 2.5D, where only one height value is permitted for each x, y position) that many geologists use. Therefore a developing part of the workflow remains the standardisation of data formats for sharing project results. The lack of commercial solutions directly providing a geological interpretation framework for lidar data has meant that in-house software has been created for more advanced data conversion, processing and interpretation tasks.

ACCURACY CONSIDERATIONS: THE UNCERTAINTY FACTOR

Despite the increased spatial accuracy and resolution offered by incorporating a technique such as lidar into the outcrop data acquisition workflow, consideration of the sources of error and the effect of these on geology data extracted is also important. Many error sources have the potential to enter the chain, which can be primary, affecting the raw data, or secondary, introduced during processing. Although the manufacturer of a laser scanner may quote a 3D point as having an accuracy of 0.01 m at 50 m range (e.g. Riegl 2007), this figure may be degraded by many factors that serve to lower the expected precision. For raw laser measurements, the point precision is likely to be degraded with greater range, poor atmospheric visibility, and the terrain type being measured. Additionally, the angle of view of the measurement is important, as a return from a surface highly oblique to the instrument is likely to contain an average of multiple ranges, especially at longer range when the laser beam width is wider (e.g. Huising & Gomes Pereira 1998). For input to most geological problems, these errors are likely to be of lesser importance, the most obvious effect being some blunders that can be removed during point cloud editing and preparation prior to triangulation.

The second type of error, resulting from processing, is more serious but often more controllable. Care should be taken to make sure that error propagation through the workflow is minimised. Error during registration will affect the alignment of the point clouds, DEM creation, and tracking features across wide areas. Editing and decimating the point cloud with unsuitable parameters may introduce too much error. During DEM creation, hole filling, interpolation and decimation can all make the resultant triangle mesh deviate from the original point data. Texturing the model with inaccurately registered imagery will make the position of interpreted features incorrect. Any of these errors in the virtual model will affect the ensuing interpretation, which will then affect surfaces and grids built from these input data. Choosing an inappropriate interpolation algorithm for building these surfaces and grids may also influence their accuracy, which will then have connotations for volume calculations and flow simulations. Uncertainty in the final models is therefore influenced by both the geology and the input geometry, a point discussed in detail by Martinius & Naess (2005).

Virtual outcrops provide spatially constrained representation of the present-day land surface. It is important to remember that this is not a 3D volume. A significant portion of the geological volume has either been removed by erosion or remains in the subsurface. Any features interpreted on either the true or virtual outcrop must be extrapolated into true three dimensions. The difference with using lidar and virtual outcrops is that the observed geometries serve to constrain the extrapolated model with higher accuracy, as the 2D slice through the subsurface is better defined than when using only traditional mapping, logging or structural measurement. Use of a quantitative and redundant data capture technique also allows quality control of the derived input to reservoir modelling software, as at all stages the extracted features can be checked with the raw point cloud data. An appreciation of the errors that may be introduced at each stage of the workflow is useful for limiting their effect, and when considering the overall accuracy of the derived products.

Laser scanning has the potential to enhance field data collection, by providing an accurate framework for capturing areas of outcrop. Creation of virtual models that can be visualised and quantitatively interrogated by the user is very useful for many projects. Use of a single project coordinate system, achieved using GPS, allows fine geometric detail to be resolved over the extent of the study area, even between areas that are not connected by exposure, as well as allowing integration of other field data. Because of the higher spatial resolution and precision, new applications at different geological scales can now be realised with greater efficiency. Although such techniques do not replace conventional geological fieldwork, as an understanding of the actual geology is still the most important factor in solving a research question, they do allow for a more integrated and quantitatively driven workflow to be implemented.

Lidar instrumentation is relatively easy to use, and it is simple to acquire a point cloud. However, it is important to stress that to gain meaningful geology data from lidar is not always as simple as collecting a point cloud. There are a number of important considerations in the data acquisition, processing and interpretation workflow that must be addressed. Standardisation of algorithms and formats for lidar processing and interpretation is yet to be achieved, making the preparation of outcrop models a nontrivial task where much vegetation or other objects that create noise in the raw point data exist. Although lidar data provide a much higher accuracy and resolution than traditional fieldwork, an awareness of the sources of error and uncertainty in the workflow, from data collection to reservoir modelling, is necessary. Where possible, checks with raw data can give a quantitative measure on the uncertainty at each stage, thus building confidence in the final products and therefore in the results of the actual geological problem being solved.

PALAEOWEATHERING FEATURES AND SUCCESSIVE SILICIFICATIONS

Most of the widespread and well-preserved silicification in inland Australia is surficial, without sedimentary cover. It affects both the older sedimentary formations (Palaeozoic and Mesozoic) and even the granites in the cratons. Thus, there are not many guides for dating the silicification. Geometric relationships between the silicification features and the Cenozoic stratigraphy and or present-day geomorphology have been used to define particular silicification periods (Simon-Coineon et al. 1996). However, the age of the silerctes remains disputed. In the past, silcretes were believed to be related to a palacosurfacc (the Cordillo Surface) around the basin margins coincident with the surface of the midEocene deposits (Kyre Formation) and of mid-Hoecnc to OligoMiocene age. The groundwater silcretes were first distinguished by their mineralogieal composition (opal-rich) and later by their general morphology and distribution (Thiry & Milnes 1991).

Some workers did not distinguish these silcrcte types and related both to a single silicification episode. Others recognised the different silicification processes but considered the groundwater silcretes to be younger than the pedogenic silcretes and mostly of MioPliocene age (Mutton ct al. 1978; Wopfncr 1978; Simon-Coinçon et al. 1996). Red-brown hardpuns are young weathering features and have not always been recognised as silicified materials. They generally developed in Pleistocene to Holoccnc sediments (Benbow 1982) and may be contemporaneous with groundwater silcretes (Simon-Coinçon et al. 1996). In this chapter we compile new field and laboratory data that allow us to provide a coherent interpretation of the evolution of silcretes and associated distinctive palacowcathering features in the Tertiary regolith of inland Australia.

Much of the study area is underlain by Mesozoic sediments of the Eromanga Basin, which overlies Proterozoic cratons and Palaeozoic folded belts. Cenozoic sediments are shallow and are restricted to a number of smaller basin areas and to veneers of clastic sediments mantling palaeosurfaces and infilling palaeomorphological features. The Eromanga Basin contains terrestrial and marine clastic sediments, largely deposited in a shallow epicontinental sea during a series of Early Cretaceous transgressions. After withdrawal of the sea, a long phase of weathering and erosion prevailed from the Cenomanian to the late Palaeoeene both in areas where Tertiary sediments wore subsequently deposited and in large areas that up to the present day have never been covered by younger sediments (both cratons as well as basin areas).

Downwarping produced a number of shallow continental sedimentary basins, the largest being the Lake Eyre and BiIla Kalinu basins, in which thin sheets of clastic material were deposited at various times during the Cenozoic. A wide marine basin (Eucla Basin) developed in southern Australia and was infilled with a thick sequence of carbonates. A major palaeo-drainage system developed that fringed the inland margin of the Eucla Basin. The Cenozoic sedimentary record is not complete. This, along with the unfossiIiferons and weathered nature of much of the succession, has caused much uncertainty in stratigraphic interpretation (Callen el til. 1495; Alley et al. 1999).

GEOMORPHOLOGY

The study area contains some significant landforms (Simon-Coinçon et al. 1496), as follows.

(1) High pediments of the Stuart Range, in the western part of the area, form an extensive erosion surface mantled with an alluvial cover. This slopes down from the Musgrave Range in the north and the Davenport Range in the cast towards the Eucla Basin. The palacosurfacc has resulted from a long period of weathering and erosion after late Albian retreat of the Cretaceous seas. It exhibits broad drainage patterns with wide dry valleys and is characteristically capped by a discontinuous veneer of silcrctes.

(2) Palaeochanncls are incised to depths of over 100 m in the Stuart Range palaeosurfacc and can be followed over several hundred kilometres (Alley el al. 1999). The cutting of the palacochannel system may have been initiated by

the opening of the rift valley along the southern coast following the separation of Australia from Antarctica. Thus, some palacochannel incision may date back to the Early Mesozoic. Infilling occurred in successive stages during the Tertiary.

(3) The Musgrave Range in the north, and the Davenport Range, an upfaulted inlier of Prccambrian bedrock near Lake Eyre, form the highs in these palaeolandscapes.

(4) A lowland complex surrounding Lake Eyre has active drainage features and a conspicuous erosional scarp to the west. The development of the lowland complex relates to the subsidence of Lake Eyre and most probably started during the late Miocene to early Pliocene (Simon-Coinçon et al. 1996).

Palaeowcathering Features

Deeply weathered rocks, often reaching 60 m in depth, are widespread on the Stuart Range palacosurface and in the areas of palaeoehannel systems. These bleached and weathered profiles with ferruginous mottling are the main feature of the study area. The profiles are well developed in flat-lying Cretaceous sediments but also occur in all geological units forming the palacosurface including granites on the edges of the Gawler Craton in the south, Cambrian and Proterozoic rocks on the western side of Lake Torrens, and Palaeozoic and Proterozoic sequences and granites on the southern flank of the Musgrave Range.

The weathering tentures include a white colour, a low rock density, the occurrence of alunite, iron oxide accumulations and the systematic presence of crystalline gypsum in fractures and voids. Profiles are invaded by fossil termite burrows, which cnn reach 40-50 m depth (Robertson & Scott 1987; Milnes et al. 1991). Termite burrows to such depths point to a very low groundwater table at that time. The profiles may contain layers of groundwater silcretes, and precious opal, which fills voids, including cracks or moulds of dissolved shells. In relation to the occurrence of termite burrows and gypsum, the bleaching is an enrly form of weathering.

Pedogenic silcretes formed near the landsurface. The typical profile comprises a nodular horizon at the base, overlain successively by prismatic columnar and pseudo-breccia horizons. Large illuviation features suggest a vadose groundwater environment, with intermittent and repetitive infiltrations and illuviations. No pedogenic silcretes occur on the surface of the lowland complex. They occur dominantly on the high pediments of the Stuart Range palaeosurface and its remnants, down to the aggradational and erosional palaeosurfaces of the palaeochannels towards the edges of the liuela Basin. Pedogenic silcretes appear to be restricted to formations older than the Miocene.

Groundwater silcretes developed in the bleached saprolite. They form flat-lying lenses or even irregular pods of porcellanite, jasper and vitreous quartzite, which may superimpose one another within the bleached formations. Preservation of primary structures (bedding, burrows, mottles, etc.) and fabrics of the enclosing rocks are characteristic of this form of silicification. Groundwater silcretes crop out on the erosional scarp of the Stuart Range and in shafts and deep quarries in the opal fields. It is likely that successions of groundwater silcretes developed in response to groundwater fluctuations (Thiry & Milnes 1991). Termite burrows in the bleached and weathered profiles are preserved in place by the silicification, indicating that groundwater silcretes developed after the bleaching, in response to a rising water table.

Red brown hardpans are regolith materials impregnated by iron oxides and hydroxides, and amorphous silica (Milnes et al. 1991). They are characterised by anastomosing subhorizontal fractures, which produce a pronounced but irregular layering. Clasts of various types including pedogenic silcrcte, jasper, porcellanite and saprolite represent remnants of underlying horizons progressively incorporated into the hardpan. Red-brown hardpans are generally developed in fluvial and colluvial materials of Pleistocene to Holocene age (Benbow 1982). In the Stuart Range pediments, the hardpans overlie pedogenic silcretes. Where there has been significant erosion, hardpans have developed on groundwater silcretes or directly on bleached Cretaceous saprolite.

Coober Pedy Opal Field

The Coober Pedy opal field has the greatest extent and exhibits a wide variety of sections. The diggings are localised in the bleached Cretaceous Bulldog Shale and often reach 30 m depth in open pits, and up to 50 m in underground mines. A cover of Tertiary clastic formations is present in some areas. The Larkins Folly section about 12 km away from the erosion scarp shows typical features.

The bleached Cretaceous sequence forms the main part of the sections and is generally white or pale yellow, but may be pale pink. Purple and red iron mottling and staining are common. Deep vertical features penetrate the bleached sediments and are infilled with red sandy clay. These are fossil termite burrows, often following vertical fractures in the bleached material. In places, the burrows may also be concentrated in horizontal zones that could represent palaeo-water-table levels. Alunite is also present in alteration zones along vertical joints, as well as in horizontal bands, where it tends to occur as rounded masses and beds of cream or white colour that replace the bleached Cretaceous sediments. In places, as in the Shell Patch field, about 30 km NW of Coober Pedy, some sections contain pure alunite beds, without quartz or clay mineral, that are up to 0.5 m in thickness and stretch over several tens of metres. In many places, alunite is restricted to small blebs 5-10 mm across. Alunite is found down to the deepest workings. In many places, infilled burrows can be observed cutting across the alunite, and thus the age relationships can be determined.

Higher in the sections, groundwater silcrete horizons may occur. Some sections are devoid of groundwater sileretes whereas others display up to three horizons. These are jasper (glossy lustre) or porcellanite (dull lustre) in the form of amoeboid masses arranged in horizontal layers with silica impregnation of joint infillings extending beneath the horizon. Such horizons vary from 0.2 to 1.0 m in thickness. Where two horizons are superimposed, the upper one is generally formed of dull and porous porcellanite disrupted by gypsum veins and collapsing structures, whereas the lower one is formed of shiny jasper and is well preserved. This may indicate that weathering of the upper horizon may be older than the weathering of deeper one. In many places, burrows with infillings are preserved and silicified within the suicided horizons, testifying to the development of the termite burrows prior to silicification.

Above the groundwater silcrete is a pedogenic silcrete of quartzitic composition with a nodular and/or massive basal zone overlain by a well-developed columnar facies. The pedogenic silcrete forms a regular horizon, which tends to be best developed in depressions. The significance of these is not clear, but they were probably structures of hydrological significance. The silcrete horizon is often disrupted by collapse structures and karst-like hollows.

Red-brown hurdpan occurs at the top of most of the sections and has a characteristic laminar structure. The hurdpan may contain void and channel fillings of opal, carbonate, palygorskite and barite, indicating a complex paragenesis.

Gypsum is everywhere in the sections. Vertical and horizontal cracks, up to 0.5 m wide, are filled with massive spar gypsum. Gypsum veins disrupt the primary structures, the alunite veins, the groundwater and the pedogenic sileretes and are also common in the hurdpan capping. Gypsum dissolution is obvious and is responsible for numerous collapse structures. The distribution of the gypsum veins and the collapse structures indicates successive dissolution and precipitation.

Opal, both precious and potch (without colour), is found in the bleached Cretaceous sequence and is generally localised in horizontal or subhorizontal structures, but also in vertical or oblique veins and occasionally as infillings and replacements of fossil shells and bones. It always infills voids formed by dissolution of gypsum veins or the collapse structures resulting from such dissolution.

In terms of mineralogical composition the bleached Cretaceous sequence contains an important percentage of opal-A, whereas opal-CT is restricted to the groundwater silcrete layers. Pedogenic silcrete contains only quartz. Opal-A is also a component of the red-brown hardpan layer. Kaolinite is usually dominant in the bleached Cretaceous sequence, but smectite appears towards the top of the sections and is dominant in the red-brown hardpan and in the soil material penetrating the pedogenic silcrete and the collapse structures.

Andamooka Opal Field

The Early Cretaceous sediments here are typically bleached, variably silicified and ferruginised. and are extensively invaded by gypsum. Bedding is well displayed in the sediments, and is outlined in places by layers of pebble- to cobble-sized boulders. The selected profile described here comes from the Lunatic field, about 3 km north of Andamooka township.

The Cretaceous sediments at the base of the section are bleached and display numerous termite burrows oiled with greybluish materials and scattered quartzite boulders (former dropstones). Groundwaler silicification resulted in the development of a more or less sandy bluish opalite made up of small granules (5-20 mm in diametre) cemented together by opal. A pedogenic silcrete, with a massive lower horizon and a thick columnar horizon at the top, rests immediately on the opalite. The columnar structures are particularly well developed, with large columns capped by broad and thick illuviated cappings. Quartzite boulders are included within the columnar silcrete. Some of these boulders form the base of a columnar structure resulting from spectacular illuvial layers capping the boulder. In places, thin veneers of the bluish opalite encrust the columns of the pedogenic silcrete. The section is topped by a crusty loamy soil (desert loam, natrargid) over highly gypseous clay with minor carbonate, which impregnates the pedogenic silcrete.

The bleached Cretaceous sequence contains about 30 40% of opal-A with the remainder being quartz and clay minerals. The opalite contains opal-CT as well as quartz and clay minerals. However, the pedogenic silcrete is composed only of quartz with 2-3% of titanium oxide. The clay fraction in the bleached Cretaceous sequence is dominated by kaolinite, whereas smectite and kaolinite occur in about equal abundance in the silicified horizons.

There are two main points of interest. First, the presence of quartzitic clasts in the pedogenic silcrete horizon suggests that the pedogenic silcrete formed in Cretaceous sediments without major disruption of the position of the dropstones. Second, the granules within the opalite are the granules and nodules from the lower horizon of the pedogenic silcrete. Cementation of these granules by the opalite and the presence of bluish opalite encrusting the columns of the pedogenic silcrete indicate that the opalite post-dates the development of the pedogenic silcrete. However, if the pedogenic silcrete developed in a vadose environment, as indicated by the large illuviation structures, and the opalite formed in a groundwater environment, the groundwater table rose between the development of the pedogenic silcrete and the formation of the opalite.

Stuart Creek Opal Field

These sections have already been described in detail (Thiry & Milnes 1991). The Cretaceous sediments here are not bleached and are unconformably overlain by a veneer of Tertiary fluviatile sediments of early to mid-Miocene age (Mirikata Formation).

The sections intersect an erosion channel cut in the Cretaceous sequence and filled with Tertiary sands, to which the silicification is largely confined. At the base of the channel, the Cretaceous sequence is essentially grey shale with gypsum-oiled joints. A massive lens of variably silicified sand, about 2 m thick and 50 m long, occurs at the base of the Tertiary infill. It consists of lustreous, glassy, varicoloured grey, beige, red and purple quartzite patches rimmed by porous sandstone. Bedding is preserved. Higher in the Tertiary sand infill there is a second silicified horizon composed of glassy quartzite with complex internal structures and numerous hollows, which are interpreted to be dissolution structures. At the top of the section is a pedogenic silcrete with elongate, centimetre-sized silica nodules at the base succeeded by a continuous unit of lateral to vertical columns.

At the base of the erosion channel, the silicification, as a pink and purple opalite, extends to a depth of 2.3 m into the underlying Cretaceous sequence. This opalite is commonly disrupted by veins of satin spar and rosette gypsum.

Fine sedimentary laminations and bioturbation features, typical of Cretaceous sediments, are conserved. The overlying beige Tertiary sand shows only irregular silicified sand lenses (about 0.5 m thick), which have a glassy quartzite appearance with pink, grey and greenish colours outlining successive stages of dissolution and silicification.

Secondary silica occurs as quartz and opal-CT. Opal-CT is restricted to opalite formed by silicification of the Cretaceous sediments, and to the lowest silicified pan in the Tertiary sequence, and is absent from the higher parts of the section, where quartz is the only secondary silica mineral. There is no Opal-A in the sequence. Alunite has been collected in veins towards the top of the Cretaceous sequence in a neighbouring section.

The Stuart Creek sections clearly show different kinds of facies generated by a single phase of groundwater silicification, specifically the development of glassy quartzite within the Tertiary sands and massive opalite within the Cretaceous shales. Superimposed silcretes, exemplified by the opalite horizons and the upper and lower silicified pans in the Tertiary sequence, are the consequence of lowering of the groundwater table following landscape dissection. Dissolution features in the upper silicified pan resulted from dissolution of earlier silica deposits by more dilute percolating water in the unsaturated zone after the water table was lowered.

Mintabie Opal Field

The Mintabie opal field is restricted to the Mintabie Beds, which are of Ordovician age and are unconformably overlain by a thin sequence of Cretaceous shales and sandstones and, in places, by Tertiary sandstones. The Mintabie Beds are a sequence of well-sorted white sandstones, with minor claystone interbeds, and crop out extensively along a scarp facing the Lake Eyre depression.

The Palaeozoic sandstones are white and soft and obviously weathered, but primary sedimentary structures are still preserved. Fossil termite burrows are present throughout the sandstones, most often following vertical joints and in places concentrating in more clayey horizontal joints. Burrows and open joints are infilled with greyish and, in some cases, red sandy materials. Yellow alunite nodules also occur in these weathered sandstones.

Tightly cemented zones, up to several metres and even several tens of metres wide, have developed within the white sandstones. These hard, cemented sandstones also contain termite burrows like those in the soft sandstone, pointing to secondary cementation of the sandstones after weathering and bioturbation.

A veneer of Cretaceous sediments is in places preserved above the white sandstone and silicified to jasper and porcellanite. Weathering structures, such as fragmentation into prismatic structures and termite burrows, are well preserved in jasper. which represents massive groundwater silicification up to 5 m thick in some sections. The upper part of the sandstone often displays karst-like dissolution pipes from few centimetres to several metres in diametre.

Pedogenic silcrete occurs at the top of the sections, in places within karst-like dissolution pipes. Large illuviation structures carpet the floor of former solution cavities and cap a sandstone breccia that fills solution pipes. The matrix between the silcrete nodules and columns is irregularly cemented, sandy and grey- to red-brown in colour. The mineralogical composition of the different facies is rather simple and similar to that of the bleached Cretaceous sections. Traces of feldspars are preserved at the base of the bleached soft sandstones. The weathered sandstones contain opal-A as well as quartz and clay minerals. The tightly cemented sandstones and the jasper contain opal-CT. Pedogenic silcretes are made up only of quartz but the matrix in which the nodules and columns are embedded also contains opal-A. Alunite formed at different levels in the sections. Gypsum is present as spar crystals in some joints, but is not as widespread as in the bleached Cretaceous sequences. However, calcite is present in numerous joint and burrow infillings, especially towards the top of the sections.

Gawler Craton

In the Gawler Ranges silcretes, associated with granitic rocks, occur on the southern edge of the high pediments of the Stuart Range palaeosurface. Relatively few studies have discussed silcretes associated with granitic rocks. The Glenloth section is situated near Mt Mitchell, a remnant of an elevated palaeosurface developed in the Precambrian Glenloth Granite, a medium- to coarse-grained granite that has been extensively bleached and weathered. The area is one of scattered plateau remnants eroded by pedimentation and with escarpments and 'breakaways' exposing deep sections through the bleached and weathered bedrock.

Bleached kaolinitic saprolite is exposed at the base of the section and passes gradationally into sandstone at the top. The fabric of the weathered granite is preserved with pseudomorphs of feldspar, mica and quartz bedding perfectly recognizable. Opal veins and centimetre- to decimetre-wide 'veins' filled with quartzosc grit penetrate the granite. Burrows are visible in some of these. The upper part of the saprolite has irregular patches of cemented and hardened saprolite. The hardened zones also intersect both the opal and the grit veins. Although the granite texture is still recognizable, iron oxide coatings and veins are 'fossilised' in these hardened cores.

The granite fabric is lost higher in the profile as the feldspar grains or pseudomorphs disappear and the quartz grains become disrupted, corroded and more closely spaced, developing a grit structure. The silerete at the top of the section displays a well-developed columnar structure typical of pedogenic silcretes. In places, granite fabric is preserved within the columns, indicating that the silcretes developed either directly from the saprolite or incorporated granite boulders or pebbles.

The similarity of the mineralogical composition of this section with those in the opal fields is remarkable. Kaolinite together with opal-A is typical of the saprolite at the base of the section where feldspars and muscovite are partly preserved. Opal veins cutting across the saprolite are a mixture of opal-A and kaolinite. The tightly cemented patches of saprolite are devoid of feldspar and clay minerals, and contain opal-CT. Kaolinite is present in the unconsolidated grit facies but no opal is detected. Pedogenic silerete contains only quartz and anatase.

PETROFABRICS, MINERALOGY AND GEOCHEMISTRY

Bleached Fades

Petrofabric and mineralogy : At all field sites the various bleached facies have many mineralogical and petrographic characteristics in common. In particular, they all contain opal-A, kaolinite and altered aluminosilicates (feldspars and clay minerals). Quartz grains in the bleached profiles never exhibit dissolution features.

Kaolinite is abundant and forms on average 15-30% of the whole rock, reaching 50% in some samples. It has several habits: (1) large flat crystals with residual micaceous sheets, which obviously derive from alteration of micas; (2) elongated concertina-like kaolinite with an average size of 5 μm × 20 μm, which is clearly neogenic; (3) small scattered plates; (4) small vermicular kaolinite crystals tangled in masses of 100-300 μm diametre, which developed by alteration of feldspar grains or, in some samples, alunite nodules.

SEM investigation shows that the large kaolinite crystals are often altered. Kaolinite sheets show dissolution features and microsphere structures at the sheet edges. This alteration may progress to a complete transformation of the kaolinite sheets. Kaolinite plates have a granular aspect and ultimately packed microspheres comparable with amorphous silica microspheres can entirely replace the kaolinitc. At this stage, the kaolinite is highly depleted in alumina, testifying to the leaching of alumina and accumulation of silica. This replacement of the kaolinite by silica microspheres is similar to the structures described by Jones et al. (1966) derived from alteration of biotite with sulphuric acid and by Sornein (1980) from alteration of muscovite in a highly acidic gossan weathering profile.

Microcline is preserved in the deepest parts of the profiles in the Mintabic opal field. Higher in the profile, and even in the granite profile at Glenloth, feldspar ghosts are conspicuous from their crystalline shape or the myrmekitic quartz structure. In the upper parts of the weathering profile there are two types of feldspar alteration: (1) replacement by more or less translucent opal embedded with scattered kaolinite flakes; (2) replacement by vermicular kaolinite crystals.

SEM shows that alteration of the feldspars to opal is accompanied by dissolution cavities along the cleavages and preservation of the structure by deposits of silica microspheres. At this stage the feldspar is depleted in alumina but not completed devoid of it.

Opal forms the main part of the matrix in the bleached Cretaceous sediments in the Coober Pedy, Andamooka and Stuart Creek opal fields, and even most of the intergranular matrix in the weathered sandstones at Mintabie. It shows high refringence and a lumpy appearance, which results from its granular structure as well as from its microporosity. Translucent and isotropic in most cases, it nevertheless often shows light birefringence, which may come from the opal itself or from the enclosed kaolinite. It is not easy to make a quantitative estimation of opal-A. Nevertheless, data from a combination of optical microscopy. X-ray diffraction (XRD) and electron probe microanalysis suggest that opal constitutes between 40 and 70% of the bleached Cretaceous facies. The bleached Cretaceous sequence at Andamooka has a lower opal and a higher kaolinite content than that at Coober Pedy.

In summary, the primary aluminosilicate minerals in the various bleached facies at all sites are altered to either kaolinite or opal. The kaolinite develops from alteration of micas and feldspars, but is also precipitated from solutions in the pores of the rock. The main feature is a very high content of opal-A in all bleached facies. There is no obvious quartz dissolution or silica leaching. Alunite has developed at the base of most of the profiles.

Geochemistry

The bleached profiles show some of the features of leached weathering profiles and also characteristics of evaporitic environments, bringing together kaolinite and opal, and even sulphates. In the past they have been ascribed to two successive weathering phases, a humid phase leading to kaolinisation and a later arid phase of silicification. On the other hand, the isotopic composition of the alunite indicates crystallisation from evaporated groundwatcr (Bird et al. 1989) and that this sulphate may have resulted from drying out of the profiles after a period of active weathering. Moreover, gypsum-bearing horizons are common features in soils of inland Australia. Gypsum occurs in all geomorphological units and has been related to aeolian inputs from the extensive salt lakes in the region (Jessup 1960) or from littoral marine deposits (Bird et al. 1989).

In fact, the close and ubiquitous association of kaolinite, opal, alunite and gypsum largely precludes the possibility of two distinct weathering phases. There has not been any substantial leaching of the weathered material (K and Ca are retained in the profiles in the sulphates alunite and gypsum, and Si in the opal varieties). This is inconsistent with an early weathering under a humid climate. Moreover, opal replacing kaolinite and feldspars with preservation of the primary mineral morphologies cannot be related to secondary opal precipitation. Weathering by an acidic sulphate-rich solution seems the only likely way to produce such proliles. The formation of alunite requires highly acidic environments, with pH as low as 2.5-3.5. Such environments are also necessary to explain replacement of feldspars and kaolinite by opal. Very acidic environments may be produced in two main ways: (1) oxidation of diagenetic pyrite in sediments or primary sulphide mineralisation in the basement; (2) oxidation and hydrolysis of dissolved ferrous ions (ferrolysis).

Bleached alunite-bearing profiles are widespread across inland Australian landscapes and affect all types of rocks. In these conditions it is difficult to hypothesise that acidity may result from oxidation of pyrite: the bleached profiles are too persistent and widespread. Ferrolysis is considered to be the most likely cause of the acidity leading to development of the bleached profiles and/or alunite. It corresponds to the oxidation and hydrolysis of dissolved Fe_2+ (Brinkman 1970: Mann 1983). The reaction is $4Fe_{2+} + 6H_2O + O_2$ [arrow right] $4FeOOH + 8H^+$.

Ferrolysis may occur during entrainment and mixing of anoxic, iron-rich water, with oxygenated meteoric water. It may occur by mixing recharge water and anoxic playa water or groundwater. Mann (1983) proposed that weathering of the iron in basement rock, followed by ferrolysis near the water table, was responsible for the acidity of shallow groundwater in Western Australia. McArthur et al. (1991) showed that the brines beneath the southwestern Australian playas are anoxic and iron-rich (Fe_{2+}) and become highly acidic (pH 2.8-4.0) at their margins. Those workers suggested that acidity develops where the brines mix with oxygenated recharging meteoric water from the water table in the surrounding dunes.

Weathering processes leading to the bleached profiles have been tested by geochemical modelling (Thiry et al. 1995). The simulation model shows that sulphate-rich oxidising brines are able to alter the Cretaceous shales and to form, in a single process, kaolinite, alunite, gypsum and opal. The bleached profiles are thus likely to be related to a warm and dry climate with a high water deficit. The dry climate is confirmed by the isotopic composition of the alunite (Bird et al. 1989).

Modelling shows that this weathering process releases considerable amounts of silica, most of which has been retained in highly soluble phases (the opal varieties) that have fed the later types of silicification in these profiles. Volume changes during modelling could correlate with disruption of sedimentary structures during crystallisation of gypsum veins and, as observed in the field section, its later collapse.

Termite Burrows

The majority of the burrow infillings are composed of round and ovoid granules averaging 200 500 μ in diametre, made from the various materials from the bleached facies, including particles of sediment with large kaolinite crystals, iron oxide granules, quartz grains, rare gypsum crystals, opal matrix, and grains of microcrystalline quartz and titanium oxide typical of the pedogenic silcrete. Generally these mineral components are surrounded by a rim of iron oxides and/or oriented clay minerals. The proportion of the different compounds is variable and there does not appear to be a specific composition. The matrix packing the granules and the fragments within the individual spheroids and ovoids is composed of smaller grains of bleached sediment, quartz and clay granules, of 10-30 μm diametre, and often stained by iron oxides. The largest intergranular spaces are lined and filled with illuviated clay.

The granules are the result of termite action. Their internal structure and oriented coating, size and distribution are similar to the structures described in present-day termite mounds in African soils (Wielemaker 1984; Eschenbrenner 1988). Termites built nests (at surface or depth in soils) and dig burrows in search of water, shaping pellets with their mandibles to remove the earth. The biological origin explains the relatively narrow granulometric distribution of the gravels.

Burrows found to depths of over 30-40 m point to low groundwater tables during and/or after development of the bleached profiles. Silicified burrows in the groundwater silicification zones indicate that the termite burrows were developed before silicification. Gravels containing fragments of pedogenic silcretes indicate that these had already developed and predated the formation of the groundwater silcretes. Generally, burrow fillings are red and of an oxidised nature. However, in places (Olympic and Black Point fields at Coober Pedy), burrow fillings are grey and free of iron oxides, indicating the development of anoxic environments after the burrows had been constructed. As silicified burrows may be red or grey, the anoxic conditions prevailed before or during groundwater silicification. Some opalite levels, as in the Olympic Held at Coober Pedy, are entirely composed of gravels similar to those of the burrows, and probably relate to silicification of underground termitaries.

GROUNDWATER SILCRETES

The silicified horizons within the bleached formations always contain opal-CT not opal-A. However, because it is difficult to detect opal-A by XRD in the presence of opal-CT, it cannot be stated unequivocally that opal-CT is exclusive of opal-A. The appearance of opal-CT in the groundwater silcretes is correlated with disappearance of the clay minerals (kaolinite and micas) that were present in the former bleached materials. On the other hand, quartz grains do not show any dissolution features.

The Opal Matrix: Under the optical microscope, the opal in the silicified horizons docs not appears any different from the opal of the bleached materials. It has the same granular aspect and a beige to brown colour.

Under the SEM, the matrix of the groundwater sileretes exhibits various forms including microspheres of opal (0.05-0.1 μm in diametre), pseudohexagonal kaolinite-like plates (0.3-1.0 μm in diametre), or flaky and honeycomb structures formed of smectite-like sheets. These forms are nearly exclusively composed of silica, with very low concentrations of alumina and other cations. They are consistent with epigenesis of sedimentary clays and consistent with the observation previously made in the groundwater silcretes at Stuart Creek opal field (Thiry & Milnes 1991).

Silica Deposits: The porosity of the groundwater silcretes is always more or less scaled by silica deposits. Silica deposits may become significant especially in samples where gypsum has dissolved, with silica replacing gypsum crystals, or more importantly when gypsum dissolution led to the formation of collapse structures. Thick silica deposits may also occur in the intergranular spaces of the loosely packed gravels in the termite burrows.

Secondary silica occurs only in the form of botryoidal void and channel cutans (laminated and oriented pedogenic coatings) of uniform thickness without any geopetal fabric. These are interpreted as deposits from mobile silica solutions in a saturated groundwater environment. The deposits contain a variety of silica minerals ranging from opal-CT to cuhcdral quartz forming depositional sequences in which there is always a gradation from poorly crystallised to better crystallised forms (Thiry & Milnes 1991). The complete sequence is: brown opal, limpid opal, lussatite or length-fast chalcedony, lutecite and euhedral quartz.

Opal-CT forms spheroliths made of criss-crossed pseudohexagonal blades, or hotryoidal deposits (lussatite) composed of lightly joined pseudohexagonal plates. A number of the silica phases have resulted from recrystallisation of primary opal deposits.

Thus, pseudo-chalcedony or chessboard-like chalcedony, displaying alternate length-slow and length-fast domains, often develops within the opal deposits, punching through the micro-laminar opal deposits indicating that the chalcedony formed by recrystallisation of opal. On a smaller scale, under the SEM, fibrous quartz (pseudo-chalcedony and lutecite) is seen to be constructed of microspherulites similar to opal microspheres. This suggests that these fibrous quartz species also formed by recrystallisation of opal. Sometimes opal recrystallisation is accompanied by the development of shrinkage cracks.

Silification Processes: Two silicification processes are evident. The opal matrix results from replacement of a former clay matrix by the alteration of clay minerals and the leaching of alumina. On the other hand, silica cutans and quartz crystals have been deposited from silica solutions. The uniform thickness of the deposits points to a water-saturated environment and accumulation of silica.

The sequence of crystallisation from the most disordered silica forms (amorphous silica and varieties of opal) towards the most ordered ones (chalcedony sheaves and quartz) points to a progressive dilution of the silicifying solution. The first deposits would have resulted from highly oversaturated solutions containing other cations and impurities, whereas the later ones would have developed in 'fresh', diluted and only slightly oversaturated waters (Milliken 1979; Thiry & Millot 1987). The compositional changes in the solutions result from progressive silicification of the host rock. Initially, the solutions are in contact with the host rock, namely, the bleached sediments with altered clay minerals and large amounts of opal, and are highly oversaturated by these poorly crystallised and/or amorphous silica phases. As silicification progresses, the solutions are isolated from the host rock by the newly formed silica rims and become less concentrated, allowing development of better crystallised silica species. The last crystallisation products primarily reflect the nature of the parent solution flowing through the silicified horizons. As the chalcedony and quartz crystallised in residual voids, the parent solutions are relatively dilute and contain only low concentrations of cations other than Si.

The disposition of the groundwater silicifications as specific levels or horizons in the profiles studied indicates clearly that they developed at the interface between two environments, Gypsum present during the silicification process demonstrates that the groundwaters were supersaturated with respect to the sulphates and thus had relatively low concentrations of silica in solution. Silica cannot he supplied by these saline groundwaters to produce the void deposits. On the other hand, if silica-rich surface waters, seeping through the profiles, mix with the sulphate-rich groundwaters. then silica solubility will be significantly lowered and silica will precipitate from the infiltrating surface waters. This is the most likely mechanism of silica precipitation in these groundwater silcretes. During high rainfall periods, the 'sweet' surface waters that become supersaturated with amorphous silicate as they percolate through the bleached profiles mix with the saline groundwaters and release most of their silica. Hence, silica precipitated in the waterlogged matrix and in the voids through which the solution flowed. Silicification of termite burrows in these horizons clearly testifies that the groundwater table had risen at that time.

Pedogenic Silcretes

The pedogenic silcretes are almost exclusively formed of quartz with some 2-4% anutase. Opal is locally present at the base of the proxies, mostly restricted to small granules and illuviation structures, and is always difficult to detect in the presence of microcrystalline quartz. The quartz is composed of coarse grains inherited from the parent material cemented by microcrystalline quartz. Quartz grains have uneven and indented contours, which look like dissolution features. In fact, under the optical microscope, these irregularities correspond to wispy crystal extinctions, which point to quartz grain overgrowths by incorporation of microquartz crystals from the cement. The titanium oxide is made of tiny grains and pigments (leucoxene), either scattered within the microcrystalline quartz matrix, or concentrated in illuviation structures and rims around nodules.

Micromorphology and Quartz Crystallisation

Illuviation features are numerous and specific to the pedogenic silcretes. Their size varies from millimetre to sub-metre scale. The most spectacular are those capping the boulders inherited from the Cretaceous shales in the Andamooka opal field sections, where they may reach 0.30 m thickness. These capping structures are made of alternate coarse laminae of size graded quartz grains and thin laminae of microcrystalline quartz and titanium oxide. The successive laminae often show erosional contacts. Successive thin illuviation laminae of microcrystalline quartz and titanium oxide differ only in their titanium oxide content. The illuviation structures develop wherever solutions have percolated, including the base of the voids that occur against the walls of vertical fissures.

The successive restructuring of the siliceous material is another feature specific to and characteristic of the pedogenic silcretes. This starts with preferential dissolution of the microcrystalline quartz matrix, releasing quartz grains and titanium oxide, which accumulate at the bottom of the dissolution features or as a cortex around nodules. Recrystallisation of the microcrystalline quartz matrix follows. Successive phases of dissolution and recrystallisation lead to accumulation of better crystallised quartz in the upper pseudo-breccia horizon than in the lower nodular and columnar horizons (Thiry 1981, 1997). Quartz overgrowths, and chalcedony and euhedral quartz develop in the upper parts of the profiles. The successive dissolution, recrystallisation and accompanying illuviation result in the complex nodular aspect of these silcretes.

Significance of the Structures

The columnar features and the illuviation and capping structures are typical soil structures related to downward percolation of water and indicate that these silcretes developed near the land surface. Profiles show a mineral sequence with poorly crystallised forms of silica, microcrystalline quartz and even opal (in some profiles) dominant at the base, and the development of quartz crystals at the top. In this sequence, the successive silica minerals derive from the previous generation by in situ dissolution and recrystallisation indicating re-equilibrium of the mineral with its environment. Silica in the main comes from dissolution at the top of the profile, followed by a sequence of precipitation and re-dissolution from top to bottom. Profiles clearly register this migration of silica from top to bottom, which results in a progressive 'sinking' of the silcrete in the landscape.

The mineral sequence probably results from a progressive concentration of the downward-moving solutions. The close link between leached and confined environments does not imply, however, a strict synchronism of degradation and construction. The two systems may work in an alternating fashion, whereas periods of loss and accumulation follow each other sequentially.

Hardpans

Hardpans are distinguishable from the other suicided materials by their red colour and their laminar structure. They are restricted to low areas in the landscape and to drainage axes, and are more or less directly related to clayey deposits, which are reported to he of Pleistocene age (Benbow 1982). Generally, they overlie and penetrate the pedogenic silcrete, but they may also directly overlie the bleached profiles. The silica deposits are almost exclusively of opal; chalcedony is very scarce.

Micromorphology

The hardpans are formed of clay minerals stained with opal and iron oxides. They are built up of glaebules and gravels of 50 200 μm diametre. Glaebules are composed of a red clayey matrix containing angular quartz grains, wrapped in microlaminated clay and iron oxide cutans and the whole impregnated with silica. They contain clasts of bleached Cretaceous sediments, groundwater silcrete, pedogenic silcrete vein opal, and even precious opal.

The deposits in the voids often include mammillary deposits of clear opal, together with or alternating with more or less ironrich clay illuviations. The clay minerals are of different kinds. Smectite is the most frequent, but palygorskite occurs too. and mixture of smectite and kaolinite is common. Lastly, calcite may cement residual porosity in some thin sections. The disposition and succession of the clay and opal deposits are spectacular. Opal concretions develop preferentially at the bottom of voids and alternate with episodes of glaebule reworking and illuviation and episodes of deposition of clay cutans. In places, opal concretions are restricted to the ceilings of voids and channels forming pendants. Opal also forms illuviation laminae at the bottoms of voids and displays desiccation cracks, pointing to a primary deposit in the form of a hydrated silica gel.

Hardpan Environment

Geopetal mammillary opal deposits and clay illuviations are typical of soil environments. Alternation of silica deposition from solutions and illuviation of clay material and even reworking of granule-sized material testify to alternating periods of water seeping gently through the profiles and periods of water flushing during rainstorms. The complex interpenetration of illuviation structures indicates intermittent and repetitive infiltration events of these types. The overall laminar structure of the hardpan profiles and the extensive development of silica concretions and clay illuviation along these flat-lying structures is probably related to a predominantly lateral flow of water in a subdued and perhaps periodically flooded landscape.

Silica concretions also indicate concentration of the soil solutions, most probably during dry seasons or periods. The origin of the silica is uncertain. It may have been derived from the profile itself by alteration of clay minerals or from topographic highs in the surrounding landscape.

Pedogenic Silcretes

Pedogenic silcretes were formed in pre-existing sediments near the land surface. They occur on the oldest palaeosurfuces forming broad gently sloping pediments and scarp foot pediments (Simon-Coinçon et al. 1996). Their features suggest a vadose groundwater environment. Intermittent and repetitive infiltration, leaching and illuviation appear to have alternated with evaporation, which caused the precipitation of silica minerals. Pedogenic silcretes are thought to develop under a climatic regime with alternate dry and wet periods (Thiry 1989). Their position on broad pediments suggests lateral migration of silica solutions and that these silcretcs developed preferentially in areas between the headwaters of streams and the depositional basin.

Simon-Coinçon et al. (1996) showed that pedogenic silcretes affect only the oldest Tertiary formations. Thus, they are regarded as one of the earliest widespread weathering features across inland Australia, developed during the late Eocene to early Oligocene.

Bleaching and Deep Weathering

The deep bleached profiles were previously thought to be related to warm and wet palaeoclimatcs (Wopfner 1978). This is inconsistent with the mincralogical assemblage of kaolinite together with opal and sulphates. Geochemical arguments and modelling suggest that the bleached profiles relate to acidic environments interacting with sulphate-rich brines. Thus, the development of these profiles probably relates to dry to sub-arid climatic conditions, rather than a tropical humid climate. The dry environments are indicated by water tables at considerable depth, as shown by the depth penetration and extent of termite burrows (Robertson & Scott 1987), the ubiquitous occurrence of gypsum that has recrystallizcd in fractures and voids (Rayot 1994), and the hydrogen isotopic composition of alunite, reflecting its formation in evaporative solutions (Bird et al. 19S9). At the present day, groundwater tables are at low levels in the landscape and sulphate-rich. It is possible that acidic alteration leading to bleaching is still active, as shown around the playas in southwestern Australia (McArthur el al. 1991).

Fossil termite burrows intersect alunite pods and clearly postdate the formation of alunite. They are infilled with red and grey soil materials, including fragments originating from the overlying pedogenic silcrctes and even red-brown hardpans. Thus, bleaching appears to be a relatively late stage of weathering in the landscape, following the development of the pedogenic silcretes, and is related to a dry palaeoclimate, saline groundwater and a low groundwater table, which may relate to incision of the landscape. K-Ar dating of alunite from Coobcr Pedy and Andamooka bleached profiles gave Miocene ages from 8.4 to 17.9 Ma (Bird et al. 1990). This is concordant with climatic conditions proposed for the Miocene (Alley et al. 1999). However, similar processes may have occurred several times during the Cenozoic in inland Australia and the bleached profiles may relate to weathering processes integrated over a long period of geological time.

8

Fractals in Geomorphology: Remote Sensing and GIS

Mandelbrot's fractal geometry is a revolution in topological space theory and, for the first time, provides the possibility of simulating and describing landscapes precisely by using a mathematical model. Fractal analysis appears to capture some "new" information that traditional parameters do not contain. A landscape should be (or is at most) statistically self-similar or statistically self-affine if it possesses a fractal nature. Mandelbrot's fractional Brownian motion (fBm) is the most useful mathematical model for simulating landscape surfaces. The fractal dimensions for different landscapes and calculated by different methods are difficult to compare. The limited size of the regions surveyed and the spatial resolution of the digital elevation models (DEMs) limit the precision and stability of the computed fractal dimension.

Fractal aspects of complex nonlinear dynamic systems are ubiquitous in the landscape and in its studied phenomena. Many natural features of the landscape have the appearance of a fractal; an example may be drainage patterns and valley networks or coast lines. Methods of fractal geometry have a mathematical basis which can be successfully applied in geomorphology. The behavior of complex natural phenomena, such as drainage systems, is at the forefront of research. Fractal dimensions and other fractal parameters in geomorphology are mainly used to quantitatively describe the topography of landscape fractal shapes and to build models of their development (Xu et al. 1993; Baas 2002).

Geomorphology is the study of the landforms and landscapes, their processes, form and sediments at the inland or coastal surface of the earth. Geomorphological mapping is the main process for providing data for the analysis of landforms and the management of land and water resources. Geomorphology constitutes the most crucial parameter in understanding Earth's surface processes, relief configuration, and landscape evolution. Furthermore, geomorphological maps and geomorphological process delineation are essential, for several other sectors of environmental research, land and water conservation organizers, natural hazard and risk managers, urban planners and construction engineers, and scientists dealing with landscapes and landforms, as well as inland and coastal land use/cover changes.

Recent advances in remote sensing, geographic information systems (GIS), and availability of a growing number of new sensors (UAV, airborne, and spaceborne) and remote-sensing-based digital elevation models have led to a revolution in the field of geomorphological mapping and enhanced the ability to understand the surface processes more clearly. Innovative remote sensing data are providing data on landform distribution, surface composition, land and water processes, inland and coastal changes, natural disasters with higher spectral, temporal, and spatial resolution. These advanced tools in addition to the extended capabilities of GIS and geospatial analysis considerably expand the capacity of geomorphological mapping.

APPLICABLE TO REMOTE SENSING

The Daubert standards reviewed above will only be applied to remotely-sensed data presented through expert testimony. Remote sensing evidence will be subject to several FRE, which are applicable whether or not an expert is called to testify.

Relevancy, Authentication, and Foundation Any evidence, scientific or otherwise, must be found relevant to the case, meaning that it must make a consequential fact more or less probable than would be deemed otherwise. If used to aid witness testimony, the map must help the trier of fact understand the testimony. Once evidence is found to be relevant, it must be authenticated. Extrinsic authentication is necessary unless the map fulfills one of the self-authentication exceptions listed in Rule 902 of the FRE. A map published by the government, for instance, is self-authenticating under Rule 902(5).

Finally, the evidence must have an adequate foundation; it must be accurate and reliable. If accuracy cannot be confirmed, courts will not admit the evidence. Of these provisions, the main evidentiary hurdle for digital maps is reliability. Courts will ask where the information in the map originated, how the information was transformed into digital form, and how the map itself was created. Since computers create digital maps, the maps will face reliability challenges as computer evidence. Courts, for instance, will inquire into "computer programming errors, equipment malfunction, data entry errors, and the volume of electronic data." Courts will also closely consider the authenticity of digital maps, particularly where the map does not meet one of the aforementioned self-authentication exceptions. As such, courts will follow Rule 901(a), requiring proof that the evidence is what its proponent claims it to be. According to Rule 901(b)(9), parties must prove that evidence encompassing a process or system, such as maps depicting remotely-sensed data, must produce an accurate result. To satisfy these rules, the experts who collected the remotely-sensed data should describe how the process operates and their involvement. Experts should also reference the data to ground information ('ground-truthing'), aerial photographs, and other maps. Logs and records of the progression from collection to presentation of the data would also verify authenticity. Technologies including steganography and cyclic redundant checksum are continually being developed to assist in ensuring the authenticity of digital imagery.

Hearsay Issues

If a map, chart, or other media is admitted to make an assertion, the evidence may be objected to on hearsay grounds. For example, remotely-sensed data could be used to create a map depicting high levels of pollution in a stream adjacent to the defendant's property. If the map is admitted to assert that the defendant caused such pollution, it may meet with a hearsay objection. If the evidence is found to be hearsay, it will only be admissible if it can be categorized as an exception to the hearsay rules. For example, Rules 803(6) and 803(8) will allow the admission of hearsay evidence that was generated by computer for use as a business or public record.

Data Characterization

A final set of rules that may pertain to the use of remotely-sensed data involve the presentation of the evidence in the courtroom. Rule 1006 allows the admission of charts, summaries, and calculations that depict a body of data too voluminous to itself be admitted into evidence for practical reasons. To avoid potential problems with admission under this rule, experts should testify that the data was correctly translated into these summary forms. If the evidence is admitted without the verification of expert testimony, Rule 1002 requires that the underlying data be admissible. For example, if a chart includes data derived from satellite photos, courts or opposing attorneys could bar the admission of the chart if the original photos do not also meet the standards of admissibility.

Constitutional Hurdles

Besides Daubert and the FRE, the Constitution presents another obstacle that remote sensing data must overcome for admission into federal courts. The main constitutional issues facing remote sensing data are allegations of invasions of privacy and warrantless searches. The Fourth Amendment states that "the right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated." Two Supreme Court cases, Dow Chemical Co. v. United States and Kyllo

v. United States, address the application of the Constitution to remote sensing data.

In Dow Chemical Co., the Court held that enhanced aerial photographs of an industrial facility taken by the EPA were admissible under the Fourth Amendment. The Court found that though commercial areas receive constitutional privacy protection, this protection does not extend to the outdoor areas of industrial complexes.

The Court also found that homes and their outside areas receive a higher level of protection than commercial areas. Still, in dicta the Court stated, "surveillance of private property by using highly sophisticated surveillance equipment not generally available to the public, such as satellite technology, might be constitutionally proscribed absent a warrant." The Court feared that technology providing information not available to the naked eye would reveal intimate details, for example, imaging that could reveal actions occurring inside a building (e.g., conversations behind closed doors or people transporting documents). Despite this concern, the Court noted that photos enhancing human vision were still admissible, provided that they do not reveal such intimate details.

The Supreme Court's latest decision regarding remote sensing data's privacy and search issues is Kyllo v. United States. Kyllo involved a police officer who used a thermal imaging device to detect heat emissions from a suspect's home. Declaring this search unconstitutional, the Court held that when "the Government uses a device that is not in general public use, to explore details of the home that would previously have been unknowable without physical intrusion, the surveillance" is unconstitutional. As in Dow Chemical Co., the Court emphasized that homes receive a high level of privacy protection under the Constitution. The Court held that, in the home, "all details are intimate details," strongly indicating that any information obtained by remote sensing data from a home's interior without a warrant would be inadmissible.

The Court did not define "general use" technology in either Dow Chemical Co. or Kyllo. Lower courts are left to speculate on what level of use might rise to this standard. For example, remote sensors that track wetland deterioration might be deemed "general use" technology if they are routinely used by the government, or if the public accepted their use. But if the device determined that someone illegally filled in a wetland in his or her backyard, that information could be inadmissible. The main lesson that can clearly be drawn from Dow Chemical Co. and Kyllo is that, in the absence of a warrant, remote sensing data will only gain courtroom admission if it does not include intimate details of commercial activity or any details from private homes.

An Overview of Remote Sensing

Remote sensing is the examination or the gathering of information about a place from a distance. Such examination can occur with devices (*e.g.*-cameras) based on the ground, and/or sensors or cameras based on ships, aircraft, satellites, or other spacecraft. Today, the data obtained is usually stored and manipulated using computers. The most common software used in remote sensing is ERDAS Imagine, ESRI, MapInfo, and ERMapper.

Modern remote sensing began in 1858 when Gaspard-Felix Tournachon first took aerial photographs of Paris from a hot air balloon. Remote sensing continued to grow from there; one of the first planned uses of remote sensing occurred during the U.S. Civil War when messenger pigeons, kites, and unmanned balloons were flown over enemy territory with cameras attached to them. The first governmental-organized air photography missions were developed for military surveillance during World Wars I and II but reached a climax during the Cold War.

Today, small remote sensors or cameras are used by law enforcement and the military in both manned and unmanned platforms to gain information about an area. Today's remote sensing imaging also includes infra-red, conventional air photos, and Doppler radar. In addition to these tools, satellites were developed during the late 20th century and are still used today to gain information on a global scale and even information about other planets in the solar system. For example, the Magellan probe is a satellite that has used remote sensing technologies to create topographic maps of Venus.

Types of Remote Sensing Data

The types of remote sensing data vary but each plays a significant role in the ability to analyze an area from some distance away. The first way to gather remote sensing data is through radar. Its most important uses are for air traffic control and the detection of storms or other potential disasters. In addition, Doppler radar is a common type of radar used in detecting meteorological data but is also used by law enforcement to monitor traffic and driving speeds. Other types of radar are also used to create digital models of elevation.

Another type of remote sensing data comes from lasers. These are often used in conjunction with radar altimeters on satellites to measure things like wind speeds and their direction and the direction of ocean currents. These altimeters are also useful in seafloor mapping in that they are capable of measuring bulges of water caused by gravity and the varied seafloor topography. These varied ocean heights can then be measured and analyzed to create seafloor maps. Also common in remote sensing is LIDAR-Light Detection and Ranging. This is most famously used for weapons ranging but can also be used to measure chemicals in the atmosphere and heights of objects on the ground.

Other types of remote sensing data include stereographic pairs created from multiple air photos (often used to view features in 3-D and/or make topographic maps), radiometers and photometers which collect emitted radiation common in infra-red photos, and air photo data obtained by earth-viewing satellites such as those found in the Landsat programme.

Applications of Remote Sensing

As with its varied types of data, the specific applications of remote sensing are diverse as well. However, remote sensing is mainly conducted for image processing and interpretation. Image processing allows things like air photos and satellite images to be manipulated so they fit various project uses and/or to create maps. By using image interpretation in remote sensing an area can be studied without being physically present there. The processing and interpretation of remote sensing images also has specific uses within various fields of study. In geology, for instance, remote sensing can be applied to analyze and map large, remote areas. Remote sensing interpretation also makes it easy for geologists in this case to identify an area's rock types, geomorphology, and changes from natural events such as a flood or landslide. Remote sensing is also helpful in studying vegetation types. Interpretation of remote sensing images allows physical and biogeographers, ecologists, those studying agriculture, and foresters to easily detect what vegetation is present in certain areas, its growth potential, and sometimes what conditions are conducive to its being there.

Additionally, those studying urban and other land use applications are also concerned with remote sensing because it allows them to easily pick out which land uses are present in an area. This can then be used as data in city planning applications and the study of species habitat, for example. Finally, remote sensing plays a significant role in GIS.

Its images are used as the input data for the raster-based digital elevation models (abbreviated as DEMs)-a common type of data used in GIS. The air photos taken during remote sensing applications are also used during GIS digitizing to create polygons, which are later put into shapefiles to create maps. Because of its varied applications and ability to allow users to collect, interpret, and manipulate data over large often not easily accessible and sometimes dangerous areas, remote sensing has become a useful tool for all geographers, regardless of their concentration.

PROCESS OF GATHERING INFORMATION IN REMOTE SENSING

Remote sensing refers to the process of gathering information about an object, at a distance, without touching the object itself. The most common remote sensing method that comes to most people's minds is the photographic image of an object taken with a camera. Remote sensing has evolved into much more than looking at objects with our eyes. It now includes using instruments, which can measure attributes about objects which unaided human eyes can't see or sense.

Some other definitions of Remote Sensing are: "Photogrammetry and Remote Sensing are the art, science and technology of obtaining reliable information about physical objects and the environment, through a process of recording, measuring and interpreting imagery and digital representations of energy patterns derived from noncontact sensor systems".

"Remote sensing may be broadly defined as the collection of information about an object without being in physical contact with the object. Aircraft and satellites are the common platforms from which remote sensing observations are made. The term remote sensing is restricted to methods that employ electromagnetic energy as the means of detecting and measuring target characteristics".

"Remote sensing is the art and science of obtaining information from a distance, *i.e.* obtaining information about objects or phenomena without being in physical contact with them. The science of remote sensing provides the instruments and theory to understand how objects and phenomena can be detected. The art of remote sensing is in the development and use analysis techniques to generate useful information".

History

In 1858 a French photographer, Gaspaed Felix Tournachon was the first to take aerial photos from a tethered balloon. A few years later in 1861, aerial photographs became a tool for military intelligence during the civil war. Aerial photographs were also taken from cameras mounted in kites (1858), and on carrier pigeons (1903). In 1909 Wilber Wright flew the first airplane to take the first photographs in flight. The first aerial photographs used in the process of creating maps was presented in a paper in 1913, by Captain Tardivo at a meeting of the International Society for Photogrammetry. Military aerial photos were used on a large scale during World War I. The military trained hundreds of people to process and interpret aerial reconnaissance photos. The French aerial units developed 56,000 photos in four days during the Meuse-Argonne offensive in 1918. After World War I and through the 1930's, commercial aerial survey companies employed many former military personnel to process aerial photos to produce maps such as topographic maps, forest management maps, and soil maps.

World War II saw the development of colour-infrared film for the US Army in 1942. These images were used to detect enemy forces and equipment that were camouflaged. A majority of Allied intelligence gathered about the enemy during this war was the direct result of aerial photoreconnaissance.

The U.S. military and other government agencies such as National Aeronautics and Space Administration (NASA) continued to develop the use of remote sensing during the cold war years. The 1960's also saw the expansion and development of earth remote sensing from space. The first military space photo reconnaissance satellite, Corona, was launched in 1960. Corona took pictures of the Soviet Union and its allies using photographic film. The exposed film was then transferred into unmanned recovery vehicles in space. The recovery vehicles then de-orbited and returned to earth by parachute carrying the film, which was then processed and analyzed in the lab. The first series of weather satellites called the Television Infrared Observation Satellites (TIROS) began launching in 1960. NASA continued collecting images for its earth observation surveys, from outer space, with the Apollo and Gemini spacecraft.

Scores of U.S. meteorological and earth observation satellites were launched during the 1970's. Also during the 1970's manned spacecraft such as the Skylab space station collected images of earth from outer space. In 1972 Landsat-1, with an original resolution of only 80 meters was the first satellite launched into space for nonmilitary earth resource observation. Landsat contained sensors capable of taking multispectral digital images. U.S. military photoreconnaissance satellites have been keep secrete and unavailable to the general public. Starting in 1976 the U.S. military started deploying more sophisticated high-resolution satellites capable of relaying digital images to earth. Eight Keyhole-11 satellites were launched between 1976 and 1988. Three improved Keyhole-11B satellites were launched between 1992 and 1996. They are able to produce images with estimated resolutions of nearly ten centimeters (four inches).

Nonmilitary satellite images have been used to monitor the degradation and pollution of the environment. These images also can be used to assess the damage of floods and natural disasters, assist in forecasting the weather, locate minerals and oil reserves, locate fish stocks, monitor ocean currents, assist in land use mapping and planning, produce geologic maps, and monitor range, forestry and agricultural resources.

Principles of Remote Sensing

We perceive the surrounding world through our five senses. Some senses (touch and taste) require contact of our sensing organs with the objects. However, we acquire much information about our surrounding through the senses of sight and hearing which do not require close contact between the sensing organs and the external objects. In another word, we are performing Remote Sensing all the time.

Generally, Remote sensing refers to the activities of recording/observing/perceiving (sensing) objects or events at far away (remote) places. In remote sensing, the sensors are not in direct contact with the objects or events being observed. The information needs a physical carrier to travel from the objects/events to the sensors through an intervening medium. The electromagnetic radiation is normally used as an information carrier in remote sensing. The output of a remote sensing system is usually an image representing the scene being observed. A further step of image analysis and interpretation is required in order to extract useful information from the image. The human visual system is an example of a remote sensing system in this general sense.

In a more restricted sense, remote sensing usually refers to the technology of acquiring information about the earth's surface (land and ocean) and atmosphere using sensors onboard airborne (aircraft, balloons) or spaceborne (satellites, space shuttles) platforms.

Satellite Remote Sensing

In this CD, you will see many remote sensing images around Asia acquired by earth observation satellites. These remote sensing satellites are equipped with sensors looking down to the earth. They are the "eyes in the sky" constantly observing the earth as they go round in predictable orbits.

Effects of Atmosphere

In satellite remote sensing of the earth, the sensors are looking through a layer of atmosphere separating the sensors from the Earth's surface being observed. Hence, it is essential to understand the effects of atmosphere on the electromagnetic radiation travelling from the Earth to the sensor through the atmosphere. The atmospheric constituents cause wavelength dependent absorption and scattering of radiation. These effects degrade the quality of images. Some of the atmospheric effects can be corrected before the images are subjected to further analysis and interpretation.

A consequence of atmospheric absorption is that certain wavelength bands in the electromagnetic spectrum are strongly absorbed and effectively blocked by the atmosphere. The wavelength regions in the electromagnetic spectrum usable for remote sensing are determined by their ability to penetrate atmosphere. These regions are known as the atmospheric transmission windows. Remote sensing systems are often designed to operate within one or more of the atmospheric windows. These windows exist in the microwave region, some wavelength bands in the infrared, the entire visible region and part of the near ultraviolet regions. Although the atmosphere is practically transparent to x-rays and gamma rays, these radiations are not normally used in remote sensing of the earth.

Optical and Infrared Remote Sensing

In Optical Remote Sensing, optical sensors detect solar radiation reflected or scattered from the earth, forming images resembling photographs taken by a camera high up in space. The wavelength region usually extends from the visible and near infrared (commonly abbreviated as VNIR) to the short-wave infrared (SWIR). Different materials such as water, soil, vegetation, buildings and roads reflect visible and infrared light in different ways. They have different colours and brightness when seen under the sun. The interpretation of optical images require the knowledge of the spectral reflectance signatures of the various materials (natural or man-made) covering the surface of the earth.

There are also infrared sensors measuring the thermal infrared radiation emitted from the earth, from which the land or sea surface temperature can be derived.

Microwave Remote Sensing

There are some remote sensing satellites which carry passive or active microwave sensors. The active sensors emit pulses of microwave radiation to illuminate the areas to be imaged. Images of the earth surface are formed by measuring the microwave energy scattered by the ground or sea back to the sensors. These satellites carry their own "flashlight" emitting microwaves to illuminate their targets. The images can thus be acquired day and night. Microwaves have an additional advantage as they can penetrate clouds. Images can be acquired even when there are clouds covering the earth surface.

A microwave imaging system which can produce high resolution image of the Earth is the synthetic aperture radar (SAR). The intensity in a SAR image depends on the amount of microwave backscattered by the target and received by the SAR antenna. Since the physical mechanisms responsible for this backscatter is different for microwave, compared to visible/infrared radiation, the interpretation of SAR images requires the knowledge of how microwaves interact with the targets.

Remote Sensing Images

Remote sensing images are normally in the form of digital images. In order to extract useful information from the images, image processing techniques may be employed to enhance the image to help visual interpretation, and to correct or restore the image if the image has been subjected to geometric distortion, blurring or degradation by other factors. There are many image analysis techniques available and the methods used depend on the requirements of the specific problem concerned. In many cases, image segmentation and classification algorithms are used to delineate different areas in an image into thematic classes. The resulting product is a thematic map of the study area. This thematic map can be combined with other databases of the test area for further analysis and utilization.

Visual System

Passive Remote Sensing: The eyes passively senses the radiation reflected or emitted from the object. The sensing system depends on an external source of illumination.

The human visual system is an example of a remote sensing system in the general sense. The sensors in this example are the two types of photosensitive cells, known as the cones and the rods, at the retina of the eyes. The cones are responsible for colour vision. There are three types of cones, each being sensitive to one of the red, green, and blue regions of the visible spectrum. Thus, it is not coincidental that the modern computer display monitors make use of the same three primary colours to generate a multitude of colours for displaying colour images. The cones are insensitive under low light illumination condition, when their jobs are taken over by the rods. The rods are sensitive only to the total light intensity. Hence, everything appears in shades of grey when there is insufficient light. As the objects/events being observed are located far away from the eyes, the information needs a carrier to travel from the object to the eyes. In this case, the information carrier is the visible light, a part of the electromagnetic spectrum. The objects reflect/scatter the ambient light falling onto them.

Part of the scattered light is intercepted by the eyes, forming an image on the retina after passing through the optical system of the eyes. The signals generated at the retina are carried via the nerve fibres to the brain, the central processing unit (CPU) of the visual system. These signals are processed and interpreted at the brain, with the aid of previous experiences. When operating in this mode, the visual system is an example of a "Passive Remote Sensing" system which depends on an external source of energy to operate. We all know that this system won't work in darkness. However, we can still see at night if we provide our own source of illumination by carrying a flashlight and shining the beam towards the object we want to observe. In this case, we are performing "Active Remote Sensing", by supplying our own source of energy for illuminating the objects.

The Planet Earth

The planet Earth is the third planet in the solar system located at a mean distance of about 1.50×10^8 km from the sun, with a mass of 5.97×10^{24} kg. Descriptions of the shape of the earth have evolved from the flat-earth model, spherical model to the currently accepted ellipsoidal model derived from accurate ground surveying and satellite measurements. A number of reference ellipsoids have been defined for use in identifying the three dimensional coordinates (*i.e.* position in space) of a point on or above the earth surface for the purpose of surveying, mapping and navigation. The reference ellipsoid in the World Geodetic System 1984 (WGS-84) commonly used in satellite Global Positioning System (GPS) has the following parameters:

- Equatorial Radius = 6378.1370 km
- Polar Radius = 6356.7523 km

The earth's crust is the outermost layer of the earth's land surface. About 29.1% of the earth's crust area is above sea level. The rest is covered by water. A layer of gaseous atmosphere envelopes the earth's surface.

The Earth's Atmosphere

The earth's surface is covered by a layer of atmosphere consisting of a mixture of gases and other solid and liquid particles. The gaseous materials extend to several hundred kilometers in altitude, though there is no well defined boundary for the upper limit of the atmosphere. The first 80 km of the atmosphere contains more than 99% of the total mass of the earth's atmosphere.

Vertical Structure of the Atmosphere

The vertical profile of the atmosphere is divided into four layers: troposphere, stratosphere, mesosphere and thermosphere. The tops of these layers are known as the tropopause, stratopause, mesopause and thermopause, respectively.

- Troposphere: This layer is characterized by a decrease in temperature with respect to height, at a rate of about 6.5°C per kilometer, up to a height of about 10 km. All the weather activities (water vapour, clouds, precipitation) are confined to this layer. A layer of aerosol particles normally exists near to the earth surface. The aerosol concentration decreases nearly exponentially with height, with a characteristic height of about 2 km.
- Stratosphere: The temperature at the lower 20 km of the stratosphere is approximately constant, after which the temperature increases with height, up to an altitude of about 50 km. Ozone exists mainly at the stratopause. The troposphere and the stratosphere together account for more than 99% of the total mass of the atmosphere.
- Mesosphere: The temperature decreases in this layer from an altitude of about 50 km to 85 km.
- Thermosphere: This layer extends from about 85 km upward to several hundred kilometers. The temperature may range from 500 K to 2000 K. The gases exist mainly in the form of thin plasma, *i.e.* they are ionized due to bombardment by solar ultraviolet radiation and energetic cosmic rays.

The term upper atmosphere usually refers to the region of the atmosphere above the troposphere.

Many remote sensing satellites follow the near polar sun-synchronous orbits at a height around 800 km, which is well above the thermopause.

Atmospheric Constituents

The atmosphere consists of the following components:

- Permanent Gases: They are gases present in nearly constant concentration, with little spatial variation. About 78% by volume of the atmosphere is nitrogen while the life-sustaining oxygen occupies 21%. The remaining one percent consists of the inert gases, carbon dioxide and other gases.
- Gases with Variable Concentration: The concentration of these gases may vary greatly over space and time. They consist of water vapour, ozone, nitrogeneous and sulphurous compounds.
- Solid and liquid particulates: Other than the gases, the atmosphere also contains solid and liquid particles such as aerosols, water droplets and ice crystals. These particles may congregate to form clouds and haze.

Electromagnetic Waves

Electromagnetic waves are energy transported through space in the form of periodic disturbances of electric and magnetic fields. All electromagnetic waves travel through space at the same speed, $c = 2.99792458 \times 10^8$ m/s, commonly known as the speed of light. An electromagnetic wave is characterized by a frequency and a wavelength. These two quantities are related to the speed of light by the equation:

Speed of light = frequency x wavelength

The frequency (and hence, the wavelength) of an electromagnetic wave depends on its source. There is a wide range of frequency encountered in our physical world, ranging from the low frequency of the electric waves generated by the power transmission lines to the very high frequency of the gamma rays originating from the atomic nuclei. This wide frequency range of electromagnetic waves constitute the Electromagnetic Spectrum.

The Electromagnetic Spectrum

The electromagnetic spectrum can be divided into several wavelength (frequency) regions, among which only a narrow band from about 400 to 700 nm is visible to the human eyes. Note that there is no sharp boundary between these regions. The boundaries shown in the above figures are approximate and there are overlaps between two adjacent regions.

Wavelength units: 1 mm = 1000 μm; 1 μm = 1000 nm.

- Radio Waves: 10 cm to 10 km wavelength.
- Microwaves: 1 mm to 1 m wavelength. The microwaves are further divided into different frequency (wavelength) bands: (1 GHz = 10^9 Hz)
  - P band: 0.3-1 GHz (30-100 cm)
  - L band: 1-2 GHz (15-30 cm)
  - S band: 2-4 GHz (7.5-15 cm)
  - C band: 4-8 GHz (3.8-7.5 cm)
  - X band: 8-12.5 GHz (2.4-3.8 cm)
  - Ku band: 12.5-18 GHz (1.7-2.4 cm)
  - K band: 18-26.5 GHz (1.1-1.7 cm)
  - Ka band: 26.5-40 GHz (0.75-1.1 cm)
- Infrared: 0.7 to 300 μm wavelength. This region is further divided into the following bands:
  - Near Infrared (NIR): 0.7 to 1.5 μm.
  - Short Wavelength Infrared (SWIR): 1.5 to 3 μm.
  - Mid Wavelength Infrared (MWIR): 3 to 8 μm.
  - Long Wanelength Infrared (LWIR): 8 to 15 μm.
  - Far Infrared (FIR): longer than 15 μm.

The NIR and SWIR are also known as the Reflected Infrared, referring to the main infrared component of the solar radiation reflected from the earth's surface. The MWIR and LWIR are the Thermal Infrared.

- Visible Light: This narrow band of electromagnetic radiation extends from about 400 nm (violet) to about 700 nm (red). The various colour components of the visible spectrum fall roughly within the following wavelength regions:

- Red: 610-700 nm
- Orange: 590-610 nm
- Yellow: 570-590 nm
- Green: 500-570 nm
- Blue: 450-500 nm
- Indigo: 430-450 nm
- Violet: 400-430 nm

- Ultraviolet: 3 to 400 nm
- X-Rays and Gamma Rays

Signature in Remote Sensing

The knowledge of spectral signatures is essential for exploiting the potential of remote sensing techniques. This knowledge enables one to identify and classify the objects of agricultural resources. It is also required for interpretation of all remotely sensed data, especially in agricultural resource data whether the interpretation is carried out visually or using digital techniques.

It also helps us in specifying requirements for any remote sensing mission *e.g.* which optimal wave length bands to be used or which type of sensor will be best suited for a particular task (agricultural survey).

All objects of agricultural resource on the surface of the earth have characteristic spectral signatures. The average spectral reflectance curves (or) spectral signatures for three typical earth's features; vegetation, soil and water. The spectral reflectance curves for vigorous vegetation manifests the "Peak-and valley" configuration. The valleys in the visible portion of the spectrum are indicative of pigments in plant leaves. Dips in reflectance that can be sent at wavelengths of 0.65 mm, 1.4 mm and 1.9 mm are attributable to absorption of water by leaves. The soil curves show a more regular variation of reflectance.

Factors that evidently affect soil reflectance are moisture content, soil texture, surface roughness and presence of organic matter. The water curves shows that from about 0.5 mm, reduction in reflectance with increasing wavelength, so that in the near infrared range, the reflectance of deep clear water is virtually zero (Mather, 1987) However, the spectral reflectance of water is significantly affected by the presence of dissolved and suspended organic and inorganic material and by the depth of the water body. Determinations of spectral signatures implies basic under standing of interaction of electromagnetic radiation with agricultural resources objects. This is also necessary for analysing and designing sensor systems for agricultural survey.

Sensor Systems in Remote Sensing

In remote sensing the acquisition of data is depending upon the sensor system used. Various remote sensing platforms (Aircraft, Satellite) are equipped with different sensor systems. Sensor is a device that receives electromagnetic radiation, converts it into a signal and presents it in a form suitable of obtaining information about the land or earth resource as used by an information gathering system. Sensor can be grouped, either on the basic of energy source. They are as classified.

Active sensor: An active sensor operates by emitting its own energy, which is needed to detect the various phenomena (*e.g.* RADAR, camera with a flash gun).

Passive sensor: The operation of passive sensor is dependent of the existing sources of energy, like sun (*e.g.* photographic systems, multispectral scanners).

The given sensor system of camera are in agricultural survey.

Photographic Cameras

The photographic system, having conventional camera with black and white photography, is the oldest and probably, so far, the most widely used sensor for recording information about ground object. Photographic cameras have been successfully used in aircraft platform remote sensing. In this system, the information is limited to size and shape, as the films used are sensitive only to visible region of spectrum. The response of black & white films is about 0.4-0.7 mm for infrared imagery, films with response extending up to 0.9 mm are available.

Return Beam Vidicon (RBM)

This is very similar to a television camera. In such a system, a fixed camera lens on a photosensitive semi-transparent sheet forms the ground image. This image is created on the surface as electrical change or potential. The TV cameras are the best example of high resolution, operated in space for resource survey was the RBV used in LAND SAT series. On LAND SAT I, II and III RBV cameras were used, each corresponding to a different wavelength band 0.475-0.585 mm (green), 0.580-0.690 mm (red) and 0.690-0.830 mm (near infrared). The Indian experimental remote sensing satellite, Bhaskara-I and II carried a two-band TV camera system, Multispectral imagery was produced in LAND SAT and Bhaskara by using separate camera tubes of each band and selecting the spectral band with appropriate filters.

Optical-mechanical Scanners

This imaging system has the advantage that any set of desired spectral bands can be selected with appropriate filter and detector combinations. The mostly widely used sensor in this category is the MSS on LAND SAT series. MSS has four spectral bands, covering form 0.5-to 1.1 mm region. MSS operates on the principle of scanning successive lines at right angles to the flight path by means of a rotation or oscillating optical system. The radiation levels along the lines are recorded by appropriate sensor elements. When used in the visible band, the collected light can be split by the optics and separately filtered and recorded, giving simultaneous multispectral recording from the one instrument. MSS can record in any part of ultraviolet to near IR window. They are use also in the thermal IR windows.

Radar and Microwave Sensors

The acquisition of data in microwave region has been possible since 1950s but its application to natural resources is considerably less developed, as compared to the visible and IR image interpretations. Microwave sensors have distinct advantages because they are unaffected by atmospheric conditions and are thus able to penetrated smoke, clouds, haze and snow. Under this system, Plan Position Indicator (PPI), Side Looking Air borne Radar (SLAR) and Synthetic Aperture Radar (SAR) can be grouped. These systems offer day and night as well as all weather capability and ability to penetrate a cover of vegetation.

Advance Remote Sensors

Linear Imaging and Self Scanning Sensors (LISS) are the advanced imaging systems. This type of scanning sensor are used an array of solid-state devices. The array may be made of photo-diodes, phototransistors or Charge-Coupled Devices (CCDs). In the LISS, the optics focuses a strip of terrain in the cross-track into the sensor array. The image from each detector is stored and shifted out sequentially to receive a video signal. The SPOT (Satellite Probatorie d' Observation de la Terra) and IRS (Indian Remote Sensing Satellite) series carry such solid-state sensor systems, which are also known as push-broom scanners. The IRS IC most advanced satellite, carries an improved sensor system. Besides carrying a sophisticated LISS-III camera, it has a Panchromatic camera (PAN) and a Wide Field Sensor (WiFS). The PAN has been designed to provide data with a spatial resolution of 5.8m in stereo mode, with a ground swath of 70km, whereas WiFS provides data in two spectral bands, with a spectral resolution of 188m and a ground swath of 180km.

DATA ACQUISITION TECHNIQUES IN REMOTE SENSING

Remote sensing is the acquisition of information about an object or phenomenon, without making physical contact with the object. In modern usage, the term generally refers to the use of aerial sensor technologies to detect and classify objects on Earth (both on the surface, and in the atmosphere and oceans) by means of propagated signals (*e.g.* electromagnetic radiation emitted from aircraft or satellites).

There are two main types of remote sensing: passive remote sensing and active remote sensing. Passive sensors detect natural radiation that is emitted or reflected by the object or surrounding area being observed. Reflected sunlight is the most common source of radiation measured by passive sensors. Examples of passive remote sensors include film photography, infrared, charge-coupled devices, and radiometers.

Active collection, on the other hand, emits energy in order to scan objects and areas whereupon a sensor then detects and measures the radiation that is reflected or backscattered from the target.

RADAR is an example of active remote sensing where the time delay between emission and return is measured, establishing the location, height, speed and direction of an object.

Remote sensing makes it possible to collect data on dangerous or inaccessible areas.

Remote sensing applications include monitoring deforestation in areas such as the Amazon Basin, glacial features in Arctic and Antarctic regions, and depth sounding of coastal and ocean depths.

Military collection during the cold war made use of stand-off collection of data about dangerous border areas. Remote sensing also replaces costly and slow data collection on the ground, ensuring in the process that areas or objects are not disturbed.

Orbital platforms collect and transmit data from different parts of the electromagnetic spectrum, which in conjunction with larger scale aerial or ground-based sensing and analysis, provides researchers with enough information to monitor trends such as El Niño and other natural long and short term phenomena. Other uses include different areas of the earth sciences such as natural resource management, agricultural fields such as land usage and conservation, and national security and overhead, ground-based and stand-off collection on border areas. By satellite, aircraft, spacecraft, buoy, ship, and helicopter images, data is created to analyze and compare things like vegetation rates, erosion, pollution, forestry, weather, and land use. These things can be mapped, imaged, tracked and observed. The process of remote sensing is also helpful for city planning, archaeological investigations, military observation and geomorphological surveying.

Data Acquisition Techniques

The basis for multispectral collection and analysis is that of examined areas or objects that reflect or emit radiation that stand out from surrounding areas.

Passive Remote Sensing

Passive systems collect data from energy that is reflected or radiated off the Earth's surface and atmosphere. A typical image derived from an infrared passive sensor consists of small equal areas referred to as pixels (7) arranged in regular rows and columns.

Each pixel has a numerical value called a digital number (DN) that records the intensity of electromagnetic energy measured for the area of ground represented by the pixel. The DN range from 0 to some higher number on a gray-scale. Each pixel is also given x and y coordinates to place it. The image can therefore be described in strictly numeric terms on a three-coordinate system with x and y locating the Pixel and z giving the DN displayed as a gray scale intensity value.

Passive sensors are described in terms of their spatial, spectral, and temporal resolutions. The spatial resolution of a sensor is the smallest area that is recorded as a separate unit (pixel). For instance, one-meter spatial resolution means that one pixel of a digital image represents an area on the Earth's surface measuring one meter in length by one meter in width. Spectral resolution refers to the number and dimension of bands (or wavelengths) of the electromagnetic spectrum that a sensor records. The higher the number of bands, the greater the sensor's ability to distinguish between objects. Temporal resolution, also known as repeat time, is the frequency with which a sensor passes over the same area.

Active Remote Sensing

Active remote sensing devices, on the other hand, emit high-energy electromagnetic radiation and record the relative amount and pattern of the energy that is reflected back. Many of these devices operate at wavelengths that not only penetrate cloud cover, but also vegetative cover and soil surfaces. The tradeoff for greater imaging capabilities, however, is increased complexity in data interpretation, as compared to passive sensor data interpretation.

Data Processing

After the satellite records the data, it is transmitted to a ground station for calibration and storage. The data may undergo various levels of processing before it is made available to the user. These levels range from simply correcting for transmission errors to performing advanced correction and analysis with model algorithms, depending on the needs of the scientists or user.

Once the data has undergone initial processing techniques, users may apply it for various purposes, from the simple production of an enhanced image to the more complex creation of image maps, thematic maps, and spatial databases. The data may also be used to develop statistical observations and graphs of the observed phenomena. To create maps and spatial databases, the initial data must be combined with other spatial data. An effective method to analyse the remote sensing data with reference to other spatial data is in a geographic information system (GIS).

Remote Sensing Data Integration with Geographic Information Systems (GIS)

Geographic information systems (GIS) are defined as computer systems capable of assembling, storing, manipulating, and displaying geographically referenced information (i.e. data points identified with respect to their location). GIS store information about the world as a collection of thematic layers that can be linked together by geography.

Remote sensing data applications and GIS have an established history of interdependency. GIS provides a format to distribute remote sensing data and to derive useable information from the data. Remotely-sensed data is also a critical means to create base GIS maps and update many data layers in the GIS. The integration of remotely-sensed data and GIS is particularly attractive because 1) the conversion of remotely-sensed raster-format data to GIS vector-format data is inexpensive and 2) remote sensing data offers a cost-effective way to visualize large geographic areas in a digital format.

There are two defining features of all GIS: the ability to overlay spatial data and the ability to change as new data becomes available. The first key feature of GIS programs is the capability to overlay multiple sets of databases into a map format that graphically explains the relationships between the data. Spatial data (points, boundaries, and lines) comprise the base of the map and can be supplemented with tabular data and image data (such as that from satellites). This powerful and versatile concept has proven invaluable for solving many real-world problems, from recording details of land use planning applications to modeling global atmospheric circulation cycles. The second key feature of GIS is their status as "dynamic maps" that can be updated and altered as needed. These maps may also be manipulated to perform scientific analyses and to create models of different environments.

In a simplistic example of GIS application, a map of city streets could be combined with latitude/longitude-referenced traffic flow data to create a map that reveals areas of frequent accident occurrence, potential detour routes, and even alternatives to improve traffic routing and alleviate rush hour stress. The same base map also may be reused to show, for example, changes in traffic patterns across time.

Urban Textural Analysis from Remote Sensor Data

Despite the new generation of very high spatial resolution sensor data (IKONOS from 1999 and QuickBird from 2001), predicted improvements in classification accuracy of urban land covers (and subsequent inference of urban land use) have yet to materialize substantially. Much of the obstruction to quality information extraction is still due to the traditional limitations of classifying image data representing urban areas: the high spatial arrangement of complex urban features and how to configure multispectral responses from land cover features into organized urban land-use categories (Barr, Barnsley, and Steel 2004). When launched, the desired objective of high spatial resolution sensor data was for increased clarity of terrestrial features, especially urban objects, by reducing per-pixel spectral heterogeneity and thereby improving land cover identification.

Clarity is certainly more evident in these finer-scale data than those from preceding sensors, but paradoxically this greater level of detail is also translated into many more unique per-pixel spectral combinations. For example, the residential land-use category can now be defined from much wider spectral variations, representing minute compositional mixtures of urban land covers, such as roads, houses, grasses, trees, bare soil, shrubs, and swimming pools, each conceivably a different residential land-use category. Following on, another limitation for improved information extraction from high spatial resolution sensor data is the reliance on techniques using traditional per-pixel spectral differentiation. To us this seems counterintuitive and we would like to see more neighbourhood-related methods, using textural and spatial parameters when dealing with fine-resolution image data.

Where traditional spectral approaches are designed to identify homogeneous features regardless of shape, textural and spatial algorithms measure both the variance within and the geometric configuration of whole urban objects, respectively. As a contribution to the growing literature, we outline an object-based pattern recognition technique that accommodates the concept of lacunarity for characterizing the textural properties of urban land cover (and therefore inferring land use) from high spatial resolution image data. In doing so, we consolidate the utility of geometric models not only for image data but for all discrete and textural spatial representations (Zhao and Stough 2005). Indeed, the ability to characterize the shapes of individual and groups of objects is a rapid area of research in computational geometry and at the heart of the recent developments in object-based models in many geographic information system algorithms.

Recall that remote sensor data are composed of multispectral pixel vectors that represent geographical objects and their relative configuration. We strongly adhere to the paradigm that geometric patterns, such as lacunarity, are valuable precursors for functional processes; in our application, the texture and spatial orientation of land-cover patterns derived from remote sensor data are both forerunners for analysing land-use juxtaposition and dynamic urban processes.

Lacunarity Approach

The lacunarity of an object is the counterpart of its fractal dimension. Lacunarity methods for urban analysis, and indeed many other applications in geospatial research, have already been reported by a number of researchers. Essentially, lacunarity is related to the spatial distribution of gap or hole sizes. For low-lacunarity measurements, all gap sizes are the same and geometric objects are deemed homogeneous; conversely, for high-lacunarity gap sizes are variable and objects are therefore heterogeneous. In other words, the variance or texture of gap sizes within the spatial delineation of geometric objects determines the level of lacunarity. Of course, textures that are homogeneous at small scales can be quite heterogeneous at large scales, and vice versa; therefore, lacunarity can be considered a scale-dependent measure of texture. Methods for calculating the lacunarity of objects were first given, in general terms, by Mandelbrot (1983) and were later implemented by various computer algorithms.

Work by Myint and Lam (2005b) developed two modified lacunarity algorithms: the binary approach, which was first introduced by Plotnick, Gardner, and O'Neill (1993), and a gray-scale routine, initially devised by Voss (1986) and used to test the effectiveness of lacunarity on high spatial resolution sensor data. This same desire to extract urban objects from fine-scale sensor data also forms the basis of this study, where we examine modifications to a differential box counting algorithm, first formulated by Dong (2000b).

Our study will also introduce two different gliding box approaches. The first uses overlapping boxes, in which the gliding box moves to a pixel next to the previous position, and the second uses skipping boxes, in which the gliding box skips the entire coverage of the previous box before moving to the next position.

As a background, and according to the gliding box algorithm proposed by Allain and Cloitre (1991), n(M,r) can be defined as the number of gliding boxes with radius rand mass M. The computations of lacunarity values are given by worked examples where the overlapping box method is demonstrated by a 4 x 4 image or local window, while a 6 x 6 image is used to illustrate the skipping box method. The 3 x 3 gliding box used in both is the base of the cube box and is always an odd number to allow the computed value to be assigned to a central cell.

A column with more than one cube box may be required to cover the maximum image intensity values by stacking cube boxes on top of each other.

The number of cube boxes required to cover the image intensity surface depends on the pixel values in the 3 x 3 gliding box. Intensity values using the example are calculated at the first, second, third, and fourth positions of the cube boxes (overlapping and skipping boxes).

The minimum and maximum pixel values are 7 and 18, respectively, at the first position of the gliding box. With a cube box of 3 x 3 x 3, these values fall in box number 3 (value of u) and 6 (value of v), respectively. The relative height of the column is then 6-3 + 1 = 4 (u-v + 1). In the same way, we can compute the required parameters for all positions of the cube boxes as follows: For the second position of the gliding box, u= 7, v= 1 (the relative height of the column is (u-v + 1) or 7-1 + 1 = 7). For the third position, u=6; v=5 (the relative height is 6-5 + 1 =2).

GIS THROUGH HISTORY

Some 35,000 years ago, Cro-Magnon hunters drew pictures of the animals they hunted on the walls of caves near Lascaux, France. Associated with the animal drawings are track lines and tallies thought to depict migration routes. These early records followed the two-element structure of modern geographic information systems (GIS): a graphic file linked to an attribute database.

Today, biologists use collar transmitters and satellite receivers to track the migration routes of caribou and polar bears to help design programmes to protect the animals. In a GIS, the migration routes were indicated by different colours for each month for 21 months. Researchers then used the GIS to superimpose the migration routes on maps of oil development plans to determine the potential for interference with the animals.

Mapmaking

Researchers are working to incorporate the mapmaking processes of traditional cartographers into GIS technology for the automated production of maps. One of the most common products of a GIS is a map. Maps are generally easy to make using a GIS and they are often the most effective means of communicating the results of the GIS process. Therefore, the GIS is usually a prolific producer of maps. The users of a GIS must be concerned with the quality of the maps produced because the GIS normally does not regulate common cartographic principles. One of these principles is the concept of generalization, which deals with the content and detail of information at various scales. The GIS user can change scale at the push of a button, but controlling content and detail is often not so easy. Mapmakers have long recognized that content and detail need to change as the scale of the map changes. For example, the State of New Jersey can be mapped at various scales, from the small scale of 1:500,000 to the larger scale of 1:250,000 and the yet larger scale of 1:100,000,but each scale requires an appropriate level of generalization.

Site Selection

The U.S. Geological Survey (USGS), in a cooperative project with the Connecticut Department of Natural Resources, digitized more than 40 map layers for the areas covered by the USGS Broad Brook and Ellington 7.5-minute topographic quadrangle maps. This information can be combined and manipulated in a GIS to address planning and natural resource issues. GIS information was used to locate a potential site for a new water well within half a mile of the Somers Water Company service area. To prepare the analysis, cartographers stored digital maps of the water service areas in the GIS. They used the proximity function in the GIS to draw a half-mile buffer zone around the water company service area. This buffer zone was the "window" used to view and combine the various map coverages relevant to the well site selection.

The land use and land cover map for the two areas shows that the area is partly developed. A GIS was used to select undeveloped areas from the land use and land cover map as the first step in finding well sites. The developed areas were eliminated from further consideration. The quality of water in Connecticut streams is closely monitored. Some of the streams in the study area were known to be unusable as drinking water sources. To avoid pulling water from these streams into the wells, 100-meter buffer zones were created around the unsuitable streams using the GIS, and the zones were plotted on the map. The areas in blue have the characteristics desired for a water well site.

Point sources of pollution are recorded by the Connecticut Department of Natural Resources. These records consist of a location and a text description of the pollutant. To avoid these toxic areas, a buffer zone of 500 meters was established around each point. This information was combined with the previous two map layers to produce a new map of areas suitable for well sites. Points sources of pollution in the water service area are identified and entered into a GIS.

The map of surficial geology shows the earth materials that lie above bedrock. Since the area under consideration in Connecticut is covered by glacial deposits, the surface consists largely of sand and gravel, with some glacial till and fine-grained sediments. Of these materials, sand and gravel are the most likely to store water that could be tapped with wells. Areas underlain by sand and gravel were selected from the surficial geology map. They were combined with the results of the previous selections to produce a map consisting of: (1) sites in underdeveloped areas underlain by sand and gravel, (2) more than 500 meters from point sources of pollution, and (3) more than 100 meters from unsuitable streams. A map that shows the thickness of saturated sediments was created by using the GIS to subtract the bedrock elevation from the surface elevation. For this analysis, areas having more than 40 feet of saturated sediments were selected and combined with the previous overlays. The resulting site selection areas that are undeveloped, are situated outside the buffered pollution areas, and are underlain by 40 feet or more of water-saturated sand and gravel. Because of map resolution and the limits of precision in digitizing, the very small polygons (areas) may not have all of the characteristics analysed, so another GIS function was used to screen out areas smaller than 10 acres. The final six sites are displayed with the road and stream network and selected place names for use in the field. Potential water well sites, roads, streams and place names. The process illustrated by this site selection analysis has been used for many common applications, including transportation planning and waste disposal site location. The technique is particularly useful when several physical factors must be considered and integrated over a large area.

Emergency Response Planning

A GIS was used to combine road network and earth science information to analyse the effect of an earthquake on the response time of fire and rescue squads. The area covered by the USGS Sugar House 7.5-minute topographic quadrangle map was selected for the study because it includes both undeveloped areas in the mountains and a part of Salt Lake City. Detailed earth science information was available for the entire region.

The road network from a USGS digital line graph includes information on the types of roads, which range from rough trails to divided highways. The locations of fire stations were plotted on the road network. A GIS function called network analysis was used to calculate the time necessary for emergency vehicles to travel from the fire stations to different areas of the city. The network analysis function considers two elements: (1) distance from the fire station, and (2) speed of travel based on the type of road. The analysis shows that under normal conditions, most of the area within the city will be served in less than 7 minutes and 30 seconds because of the distribution and density of fire stations and the continuous network of roads.

The accompanying illustration depicts the blockage of the road network that would result from an earthquake, assuming that any road crossing the fault trace would become impassable. The primary effect on emergency response time would occur in neighbourhoods west of the fault trace, where travel times from the fire stations would be noticeably lengthened. Road network of area covered by the Sugar House quadrangle plotted from USGS digital line graph data, indicating the locations of fire stations and travel times of emergency vehicles.

Areas in blue can receive service within 2½minutes, area in green within 5 minutes, areas in yellow within 7½ minutes, and areas in magenta within 10 minutes. Areas in white cannot receive service within 10 minutes. After faulting, initial model. Network analysis in a GIS produces a map of travel times from the stations after faulting. The fault is in red. Emergency response times have increased for areas west of the fault. The Salt Lake City area lies on lake sediments of varying thicknesses.

These sediments range from clay to sand and gravel, and most are water-saturated. In an earthquake, these materials may momentarily lose their ability to support surface structures, including roads.

The potential for this phenomenon, known as liquefaction, is shown in a composite map portraying the inferred relative stability of the land surface during an earthquake. Areas near the fault and underlain by thick, loosely consolidated, water-saturated sediments will suffer the most intense surface motion during an earthquake. Areas on the mountain front with thin surface sediments will experience less additional ground acceleration. The map of liquefaction potential was combined with the road network analysis to show the additional effect of liquefaction on response times.

The areas near the fault, as well as those underlain by thick, water-saturated sediments, are subject to more road disruptions and slower emergency response than are other areas of the city. Map of potential ground l liquefaction during an earthquake.

The least stable areas are shown by yellows and oranges, the most stable by grays and browns. After faulting, final model. An earthquake on emergency travel times is reduced by combining the liquefaction potential information.

GIS: COMPUTER SYSTEM OF CAPTURING, STORING AND ANALYSING

A GIS is a computer system capable of capturing, storing, analysing, and displaying geographically referenced information; that is, data identified according to location. Practitioners also define a GIS as including the procedures, operating personnel, and spatial data that go into the system.

The power of a GIS comes from the ability to relate different information in a spatial context and to reach a conclusion about this relationship. Most of the information we have about our world contains a location reference, placing that information at some point on the globe.

When rainfall information is collected, it is important to know where the rainfall is located. This is done by using a location reference system, such as longitude and latitude, and perhaps elevation. Comparing the rainfall information with other information, such as the location of marshes across the landscape, may show that certain marshes receive little rainfall. This fact may indicate that these marshes are likely to dry up, and this inference can help us make the most appropriate decisions about how humans should interact with the marsh. A GIS, therefore, can reveal important new information that leads to better decisionmaking.

Many computer databases that can be directly entered into a GIS are being produced by Federal, State, tribal, and local governments, private companies, academia, and nonprofit organizations. Different kinds of data in map form can be entered into a GIS. A GIS can also convert existing digital information, which may not yet be in map form, into forms it can recognize and use. For example, digital satellite images can be analysed to produce a map of digital information about land use and land cover. Likewise, census or hydrologic tabular data can be converted to a maplike form and serve as layers of thematic information in a GIS.

Data Capture

How can a GIS use the information in a map? If the data to be used are not already in digital form, that is, in a form the computer can recognize, various techniques can capture the information. Maps can be digitized by hand-tracing with a computer mouse on the screen or on a digitizing tablet to collect the coordinates of features. Electronic scanners can also convert maps to digits. Cordinates from Global Positioning System (GPS) receivers can also be uploaded into a GIS. A GIS can be used to emphasize the spatial relationships among the objects being mapped. While a computer-aided mapping system may represent a road simply as a line, a GIS may also recognize that road as the boundary between wetland and urban development between two census statistical areas.

Data capture—putting the information into the system—involves identifying the objects on the map, their absolute location on the Earth's surface, and their spatial relationships. Software tools that automatically extract features from satellite images or aerial photographs are gradually replacing what has traditionally been a time-consuming capture process. Objects are identified in a series of attribute tables—the "information" part of a GIS. Spatial relationships, such as whether features intersect or whether they are adjacent, are the key to all GIS-based analysis.

Data Integration

A GIS makes it possible to link, or integrate, information that is difficult to associate through any other means. Thus, a GIS can use combinations of mapped variables to build and analyse new variables. Data integration is the linking of information in different forms through a GIS. For example, using GIS technology, it is possible to combine agricultural records with hydrography data to determine which streams will carry certain levels of fertilizer runoff. Agricultural records can indicate how much pesticide has been applied to a parcel of land. By locating these parcels and intersecting them with streams, the GIS can be used to predict the amount of nutrient runoff in each stream. Then as streams converge, the total loads can be calculated downstream where the stream enters a lake.

Projection and Registration

A property ownership map might be at a different scale than a soils map. Map information in a GIS must be manipulated so that it registers, or fits, with information gathered from other maps. Before the digital data can be analysed, they may have to undergo other manipulations—*projection conversions,* for example—that integrate them into a GIS. Projection is a fundamental component of mapmaking. A projection is a mathematical means of transferring information from the Earth's three-dimensional, curved surface to a two-dimensional medium—paper or a computer screen. Different projections are used for different types of maps because each projection is particularly appropriate for certain uses. For example, a projection that accurately represents the shapes of the continents will distort their relative sizes.

Since much of the information in a GIS comes from existing maps, a GIS uses the processing power of the computer to transform digital information, gathered from sources with different projections, to a common projection. An elevation image classified from a satellite image of Minnesota exists in a different scale and projection than the lines on the digital file of the State and province boundaries. The elevation image has been reprojected to match the projection and scale of the State and province boundaries.

Data Structures

Can a land use map be related to a satellite image, a timely indicator of land use? Yes, but because digital data are collected and stored in different ways, the two data sources may not be entirely compatible. Therefore, a GIS must be able to convert data from one structure to another. Satellite image data that have been interpreted by a computer to produce a land use map can be "read into" the GIS in raster format. Raster data files consist of rows of uniform cells coded according to data values. An example is land cover classification. Raster files can be manipulated quickly by the computer, but they are often less detailed and may be less visually appealing than vector data files, which can approximate the appearance of more traditional hand-drafted maps. Vector digital data have been captured as points, lines (a series of point coordinates), or areas (shapes bounded by lines). An example of data typically held in a vector file would be the property boundaries for a particular housing subdivision.

Data Modelling

It is impossible to collect data over every square meter of the Earth's surface. Therefore, samples must be taken at discrete locations. A GIS can be used to depict two-and three-dimensional characteristics of the Earth's surface, subsurface, and atmosphere from points where samples have been collected. For example, a GIS can quickly generate a map with isolines that indicate the pH of soil from test points. Such a map can be thought of as a soil pH contour map. Many sophisticated methods can estimate the characteristics of surfaces from a limited number of point measurements. Two-and three-dimensional contour maps created from the surface modelling of sample points from pH measurements can be analysed together with any other map in a GIS covering the area.

The way maps and other data have been stored or filed as layers of information in a GIS makes it possible to perform complex analyses. A crosshair pointer (top) can be used to point at a location stored in a GIS. The bottom illustration depicts a computer screen containing the kind of information stored about the location—for example, the latitude, longitude, projection, coordinates, closeness to wells, sources of production, roads, and slopes of land.

Information Retrieval

What do you know about the swampy area at the end of your street? With a GIS you can "point" at a location, object, or area on the screen and retrieve recorded information about it from offscreen files. Using scanned aerial photographs as a visual guide, you can ask a GIS about the geology or hydrology of the area or even about how close a swamp is to the end of a street. This type of analysis allows you to draw conclusions about the swamp's environmental sensitivity.

Topological Modelling

Have there ever been gas stations or factories that operated next to the swamp? Were any of these uphill from and within 2 miles of the swamp? A GIS can recognize and analyse the spatial relationships among mapped phenomena. Conditions of adjacency (what is next to what), containment (what is enclosed by what), and proximity (how close something is to something else) can be determined with a GIS.

Networks

When nutrients from farmland are running off into streams, it is important to know in which direction the streams flow and which streams empty into other streams. This is done by using a linear network. It allows the computer to determine how the nutrients are transported downstream. Additional information on water volume and speed throughout the spatial network can help the GIS determine how long it will take the nutrients to travel downstream. A GIS can simulate the movement of materials along a network of lines. These illustrations show the route of pollutants through a stream system. Flow directions are indicated by arrows.

Overlay

Using maps of wetlands, slopes, streams, land use, and soils, the GIS might produce a new map layer or overlay that ranks the wetlands according to their relative sensitivity to damage from nutrient runoff.

Data Output

A critical component of a GIS is its ability to produce graphics on the screen or on paper to convey the results of analyses to the people who make decisions about resources. Wall maps, Internet-ready maps, interactive maps, and other graphics can be generated, allowing the decisionmakers to visualize and thereby understand the results of analyses or simulations of potential events.

Framework for Cooperation

The use of a GIS can encourage cooperation and communication among the organizations involved in environmental protection, planning, and resource management. The collection of data for a GIS is costly. Data collection can require very specialized computer equipment and technical expertise. Standard data formats ease the exchange of digital information among users of different systems. Standardization helps to stretch data collection funds further by allowing data sharing, and, in many cases, gives users access to data that they could not otherwise collect for economic or technical reasons. Organizations such as the University Consortium for Geographic Information Science and the Federal Geographic Data Committee seek to encourage standardization efforts.

SPATIAL ANALYSIS WITH GIS

Given the vast range of spatial analysis techniques that have been developed over the past half century, any summary or review can only cover the subject to a limited depth. This is a rapidly changing field, and GIS packages are increasingly including analytical tools as standard built-in facilities or as optional toolsets, add-ins or 'analysts'.

In many instances such facilities are provided by the original software suppliers, whilst in other cases facilities have been developed and are provided by third parties. Furthermore, many products offer software development kits, programmeming languages and language support, scripting facilities and/or special interfaces for developing one's own analytical tools or variants.

The website Geospatial Analysis and associated book/ebook attempt to provide a reasonably comprehensive guide to the subject. The impact of these myriad paths to perform spatial analysis create a new dimension to business intelligence termed "spatial intelligence" which, when delivered via intranet, democratizes access to operational sorts not usually privy to this type of information.

Three-dimensional GIS

To more realistically analyse the effect of the Earth's terrain, we use three-dimensional models within a GIS. A GIS can display the Earth in realistic, three-dimensional perspective views and animations that convey information more effectively and to wider audiences than traditional, two-dimensional, static maps.

The U.S. Forest Service was offered a land swap by a mining company seeking development rights to a mineral deposit in Arizona's Prescott National Forest. Using a GIS, the USGS and the U.S. Forest Service created perspective views of the area to depict the terrain as it would appear after mining. To assess the potential hazard of landslides both on land and underwater, the USGS generated a three-dimensional image of the San Francisco Bay area. It created the image by mosaicking eight scenes of natural colour composite Landsat 7 enhanced thematic mapper imagery on California fault data using approximately 700 digital elevation models at 1:24,000 scale.

Graphic Display Techniques

Traditional maps are abstractions of the real world; each map is a sampling of important elements portrayed on a sheet of paper with symbols to represent physical objects. People who use maps must interpret these symbols. Topographic maps show the shape of the land surface with contour lines. Graphic display techniques in GISs make relationships among map elements more visible, heightening one's ability to extract and analyse information.

Two types of data were combined in a GIS to produce a perspective view of a part of San Mateo County, Calif. The digital elevation model, consisting of surface elevations recorded on a 30-meter horizontal grid, shows high elevations as white and low elevations as black. The accompanying Landsat thematic mapper image shows a false-colour infrared image of the same area in 30-meter pixels, or picture elements and combine the two images to produce the three-dimentional image.

Visualization

Maps have traditionally been used to explore the Earth. GIS technology has enhanced the efficiency and analytical power of traditional cartography. As the scientific community recognizes the environmental consequences of human activity, GIS technology is becoming an essential tool in the effort to understand the process of global change. Map and satellite information sources can be combined in models that simulate the interactions of complex natural systems. Through a process known as visualization, a GIS can be used to produce images— not just maps, but drawings, animations, and other cartographic products. These images allow researchers to view their subjects in ways that they never could before. The images often are helpful in conveying the technical concepts of a GIS to nonscientists.

GIS Uncertainties

GIS accuracy depends upon source data, and how it is encoded to be data referenced. Land Surveyors have been able to provide a high level of positional accuracy utilizing the GPS derived positions. The high-resolution digital terrain and aerial imagery, the powerful computers, Web technology, are changing the quality, utility, and expectations of GIS to serve society on a grand scale, but nevertheless there are other source data that has an impact on the overall GIS accuracy like: paper maps that are not found to be very suitable to achieve the desired accuracy since the aging of maps affects their dimensional stability. In developing a Digital Topographic Data Base for a GIS, topographical maps are the main source of data.

Aerial photography and satellite images are extra sources for collecting data and identifying attributes which can be mapped in layers over a location facsimile of scale. The scale of a map and geographical rendering area representation type are a very important aspects since the information content depends mainly on the scale set and resulting locatability of the map's representations.

In order to digitize a map, the map has to be checked within theoretical dimensions, then scanned into a raster format, and resulting raster data has to be given a theoretical dimension by a rubber sheeting/warping technology process. Uncertainty is a significant problem in designing a GIS because spatial data tend to be used for purposes for which they were never intended. Some maps were made many decades ago, where at that time the computer industry was not even in its perspective establishments. This has led to historical reference maps without common norms. Map accuracy is a relative issue of minor importance in cartography. All maps are established for communication ends. Maps use a historically constrained technology of pen and paper to communicate a view of the world to their users. Cartographers feel little need to communicate information based on accuracy, for when the same map is digitized and input into a GIS, the mode of use often changes. The new uses extend well beyond a determined domain for which the original map was intended and designed. A quantitative analysis of maps brings accuracy issues into focus. The electronic and other equipment used to make measurements for GIS is far more precise than the machines of conventional map analysis. The truth is that all geographical data are inherently inaccurate, and these inaccuracies will propagate through GIS operations in ways that are difficult to predict, yet have goals of conveyance in mind for original design.

Accuracy Standards for 1:24000 Scales Map: 1:24,000 ± 40.00 feet This means that when we see a point or attribute on a map, its "probable" location is within a +/- 40 foot area of its rendered reference, just as to area representations and scale. A GIS can also convert existing digital information, which may not yet be in map form, into forms it can recognize, employ for its data analysis processes, and use in forming mapping output. For example, digital satellite images generated through remote sensing can be analysed to produce a map-like layer of digital information about vegetative covers on land locations. Another fairly recently developed resource for naming GIS location objects is the Getty Thesaurus of Geographic Names which is a structured vocabulary containing about 1,000,000 names and other information about places. Likewise, researched census or hydrological tabular data can be displayed in map-like form, serving as layers of thematic information for forming a GIS map.

Data Representation

GIS data represents real objects with digital data determining the mix. Real objects can be divided into two abstractions: discrete objects and continuous fields. Traditionally, there are two broad methods used to store data in a GIS for both kinds of abstractions mapping references: raster images and vector. Points, lines, and polygons are the stuff of mapped location attribute references. A new hybrid method of storing data is that of identifying point clouds, which combine three-dimensional points with RGB information at each point, returning a "3D colour image". GIS Thematic maps then are becoming more and more realistically visually descriptive of what they set out to show or determine.

Raster

A raster data type is, in essence, any type of digital image represented by reducible and enlargeable grids. Anyone who is familiar with digital photography will recognize the Raster graphics pixel as the smallest individual grid unit building block of an image, usually not readily identified as an artifact shape until an image is produced on a very large scale. A combination of the pixels making up an image colour formation scheme will compose details of an image, as is distinct from the commonly used points, lines, and polygon area location symbols of scalable vector graphics as the basis of the vector model of area attribute rendering.

While a digital image is concerned with its output blending together its grid based details as an identifiable representation of reality, in a photograph or art image transferred into a computer, the raster data type will reflect a digitized abstraction of reality dealt with by grid populating tones or objects, quantities, cojoined or open boundaries, and map relief schemas. Aerial photos are one commonly used form of raster data, with one primary purpose in mind: to display a detailed image on a map area, or for the purposes of rendering its identifiable objects by digitization. Additional raster data sets used by a GIS will contain information regarding elevation, a digital elevation model, or reflectance of a particular wavelength of light, Landsat, or other electromagnetic spectrum indicators.

Raster data type consists of rows and columns of cells, with each cell storing a single value.

Raster data can be images with each pixel containing a colour value. Additional values recorded for each cell may be a discrete value, such as land use, a continuous value, such as temperature, or a null value if no data is available. While a raster cell stores a single value, it can be extended by using raster bands to represent RGB colours, colourmaps or an extended attribute table with one row for each unique cell value.

The resolution of the raster data set is its cell width in ground units. Raster data is stored in various formats; from a standard file-based structure of TIF, JPEG, etc. to binary large object data stored directly in a relational database management system similar to other vector-based feature classes. Database storage, when properly indexed, typically allows for quicker retrieval of the raster data but can require storage of millions of significantly sized records.

Vector

In a GIS, geographical features are often expressed as vectors, by considering those features as geometrical shapes.

Different geographical features are expressed by different types of geometry:

- *Points*: Zero-dimensional points are used for geographical features that can best be expressed by a single point reference—in other words, by simple location. Examples include wells, peaks, features of interest, and trailheads. Points convey the least amount of information of these file types. Points can also be used to represent areas when displayed at a small scale. For example, cities on a map of the world might be represented by points rather than polygons. No measurements are possible with point features.
- *Lines or polylines*: One-dimensional lines or polylines are used for linear features such as rivers, roads, railroads, trails, and topographic lines. Again, as with point features,

linear features displayed at a small scale will be represented as linear features rather than as a polygon. Line features can measure distance.

- *Polygons*: Two-dimensional polygons are used for geographical features that cover a particular area of the earth's surface. Such features may include lakes, park boundaries, buildings, city boundaries, or land uses. Polygons convey the most amount of information of the file types. Polygon features can measure perimeter and area.

Each of these geometries are linked to a row in a database that describes their attributes. For example, a database that describes lakes may contain a lake's depth, water quality, pollution level.

This information can be used to make a map to describe a particular attribute of the dataset. For example, lakes could be coloured depending on level of pollution. Different geometries can also be compared. For example, the GIS could be used to identify all wells that are within one kilometre of a lake that has a high level of pollution.

Vector features can be made to respect spatial integrity through the application of topology rules such as 'polygons must not overlap'. Vector data can also be used to represent continuously varying phenomena. Contour lines and triangulated irregular networks are used to represent elevation or other continuously changing values. TINs record values at point locations, which are connected by lines to form an irregular mesh of triangles. The face of the triangles represent the terrain surface.

Advantages and Disadvantages

There are some important advantages and disadvantages to using a raster or vector data model to represent reality:

- Raster datasets record a value for all points in the area covered which may require more storage space than representing data in a vector format that can store data only where needed.
- Raster data allows easy implementation of overlay operations, which are more difficult with vector data.

- Vector data can be displayed as vector graphics used on traditional maps, whereas raster data will appear as an image that may have a blocky appearance for object boundaries.
- Vector data can be easier to register, scale, and re-project, which can simplify combining vector layers from different sources.
- Vector data is more compatible with relational database environments, where they can be part of a relational table as a normal column and processed using a multitude of operators.
- Vector file sizes are usually smaller than raster data, which can be 10 to 100 times larger than vector data.
- Vector data is simpler to update and maintain, whereas a raster image will have to be completely reproduced.
- Vector data allows much more analysis capability, especially for "networks" such as roads, power, rail, telecommunications, etc. Raster data will not have all the characteristics of the features it displays.

Non-spatial Data

Additional non-spatial data can also be stored along with the spatial data represented by the coordinates of a vector geometry or the position of a raster cell. In vector data, the additional data contains attributes of the feature. For example, a forest inventory polygon may also have an identifier value and information about tree species. In raster data the cell value can store attribute information, but it can also be used as an identifier that can relate to records in another table.

Software is currently being developed to support spatial and non-spatial decision-making, with the solutions to spatial problems being integrated with solutions to non-spatial problems. The end result with these Flexible Spatial Decision-Making Support Systems is expected to be that non-experts will be able to use GIS, along with spatial criteria, and simply integrate their non-spatial criteria to view solutions to multi-criteria problems. This system is intended to assist decision-making.

Data Capture

Data capture—entering information into the system—consumes much of the time of GIS practitioners. There are a variety of methods used to enter data into a GIS where it is stored in a digital format. Existing data printed on paper or PET film maps can be digitized or scanned to produce digital data. A digitizer produces vector data as an operator traces points, lines, and polygon boundaries from a map. Scanning a map results in raster data that could be further processed to produce vector data. Survey data can be directly entered into a GIS from digital data collection systems on survey instruments using a technique called Coordinate Geometry.

Positions from a Global Navigation Satellite System like Global Positioning System, another survey tool, can also be directly entered into a GIS. Current trend is data collection and field mapping carried out directly with field computers. New technologies allow to create maps as well as analysis directly in the field, projects are more efficient and mapping is more accurate. Remotely sensed data also plays an important role in data collection and consist of sensors attached to a platform. Sensors include cameras, digital scanners and LIDAR, while platforms usually consist of aircraft and satellites.

The majority of digital data currently comes from photo interpretation of aerial photographs. Soft copy workstations are used to digitize features directly from stereo pairs of digital photographs. These systems allow data to be captured in two and three dimensions, with elevations measured directly from a stereo pair using principles of photogrammetry. Currently, analog aerial photos are scanned before being entered into a soft copy system, but as high quality digital cameras become cheaper this step will be skipped.

Satellite remote sensing provides another important source of spatial data. Here satellites use different sensor packages to passively measure the reflectance from parts of the electromagnetic spectrum or radio waves that were sent out from an active sensor such as radar. Remote sensing collects raster data that can be further processed using different bands to identify objects and classes of interest, such as land cover.

When data is captured, the user should consider if the data should be captured with either a relative accuracy or absolute accuracy, since this could not only influence how information will be interpreted but also the cost of data capture. In addition to collecting and entering spatial data, attribute data is also entered into a GIS.

For vector data, this includes additional information about the objects represented in the system. After entering data into a GIS, the data usually requires editing, to remove errors, or further processing.

For vector data it must be made "topologically correct" before it can be used for some advanced analysis. For example, in a road network, lines must connect with nodes at an intersection. Errors such as undershoots and overshoots must also be removed. For scanned maps, blemishes on the source map may need to be removed from the resulting raster. For example, a fleck of dirt might connect two lines that should not be connected.

Raster-to-vector Translation

Data restructuring can be performed by a GIS to convert data into different formats. For example, a GIS may be used to convert a satellite image map to a vector structure by generating lines around all cells with the same classification, while determining the cell spatial relationships, such as adjacency or inclusion.

More advanced data processing can occur with image processing, a technique developed in the late 1960s by NASA and the private sector to provide contrast enhancement, false colour rendering and a variety of other techniques including use of two dimensional Fourier transforms. Since digital data is collected and stored in various ways, the two data sources may not be entirely compatible. So a GIS must be able to convert geographic data from one structure to another.

Projections, Coordinate Systems and Registration

A property ownership map and a soils map might show data at different scales. Map information in a GIS must be manipulated so that it registers, or fits, with information gathered from other maps. Before the digital data can be analysed, they may have to undergo other manipulations—projection and coordinate conversions, for example—that integrate them into a GIS.

The earth can be represented by various models, each of which may provide a different set of coordinates for any given point on the Earth's surface. The simplest model is to assume the earth is a perfect sphere. As more measurements of the earth have accumulated, the models of the earth have become more sophisticated and more accurate. In fact, there are models that apply to different areas of the earth to provide increased accuracy.

Projection is a fundamental component of map making. A projection is a mathematical means of transferring information from a model of the Earth, which represents a three-dimensional curved surface, to a two-dimensional medium—paper or a computer screen. Different projections are used for different types of maps because each projection particularly suits specific uses.

For example, a projection that accurately represents the shapes of the continents will distort their relative sizes. Since much of the information in a GIS comes from existing maps, a GIS uses the processing power of the computer to transform digital information, gathered from sources with different projections and/or different coordinate systems, to a common projection and coordinate system.

Adding the Element of Time

The condition of the Earth's surface, atmosphere, and subsurface can be examined by feeding satellite data into a GIS. GIS technology gives researchers the ability to examine the variations in Earth processes over days, months, and years. As an example, the changes in vegetation vigor through a growing season can be animated to determine when drought was most extensive in a particular region. The resulting normalized vegetation index represents a rough measure of plant health. Working with two variables over time will allow researchers to detect regional differences in the lag between a decline in rainfall and its effect on vegetation. The satellite sensor used in this analysis is the advanced very high resolution radiometer (AVHRR), which detects the amounts of energy reflected from the Earth's surface at a 1-kilometer resolution twice a day. Other sensors provide spatial resolutions of less than 1 meter.

Serving GIS over the Internet

Through Internet map server technology, spatial data can be accessed and analysed over the Internet. For example, current wildfire perimeters are displayed with a standard web browser, allowing fire managers to better respond to fires while in the field and helping homeowners to take precautionary measures.

The Future of GIS

Environmental studies, geography, geology, planning, business marketing, and other disciplines have benefitted from GIS tools and methods. Together with cartography, remote sensing, global positioning systems, photogrammetry, and geography, the GIS has evolved into a discipline with its own research base known as geographic information sciences. An active GIS market has resulted in lower costs and continual improvements in GIS hardware, software, and data. These developments will lead to a much wider application of the technology throughout government, business, and industry. GIS and related technology will help analyse large datasets, allowing a better understanding of terrestrial processes and human activities to improve economic vitality and environmental quality.

9

Applied Geomorphology : Nature and Objectives

APPLIED GEOMORPHOLOGY

Geomorphology has traditionally focused on the study of landforms and on the processes involved in their formation. Applied geomorphology is the practical application of this study to a range of environmental issues, both in terms of current problems and of future prediction. Applied geomorphology provides a strategic tool for informed decision-making in public policy development and in environmental resource management. Key areas of application include specific environmental settings, such as the coastal zone or dryland environments; the impacts of land use and management practice on Earth surface processes; and areas susceptible to natural hazards.

Over 60 per cent of the world's population live in the coastal zone in environments ranging from coral atolls, reclaimed or natural wetlands, dune-backed beaches, and barrier islands to cliff tops. Settlements under threat from coastal erosion and flooding from storm events, sea surges, and rising sea level lobby for protective engineering measures to prevent loss of property, livelihood, and life. Geomorphology has several applications in settings of this type. An understanding of coastal landforms and the processes acting upon them can be used to map areas at risk from cliff failure, beach erosion, and flooding. This approach is of interest to potential developers and the insurance industry and is an important tool in environmental impact assessment.

An understanding of the geomorphology of the coastal zone can also be used to predict the effects of modifying the coastal system. The installation of groynes, breakwaters, or protective sea walls has knock-on effects on the natural circulation of water and sediment in the near-shore environment.

Artificially stabilizing cliffs to prevent erosion may seem the obvious solution for cliff-top dwellers, but a geomorphological evaluation might predict that this approach could starve beaches of the sediment provided by natural cliff fall, with a consequent impact on longshore drift of sediment, and would relocate the focus of erosion further along the coast.

The nature of the problem may thus change from cliff failure at one site to beach erosion and subsequent flooding at another. An understanding of the nature and complexity of coastal dynamics is thus an essential component of a coastal-zone management strategy and is important in predicting the future effects on coastal landforms of a rise in sea level.

River-management strategies for flood alleviation have often adopted engineering solutions concentrated in particular river reaches, which are usually in areas of urban development. Reach-specific intervention measures include lining the natural channel with concrete to prevent erosion and bank instability, channel straightening to force flood water to flow rapidly through particular reaches, and flow-control structures such as sluice gates and reservoirs to control water level. These artificial measures are not always successful in preventing flooding and erosion within the river catchment, and natural sections further downstream may be overwhelmed by the river at peak flood.

The engineered reaches of rivers often become a sterile landscape because fast-flowing water in a concrete-lined channel, with minimal variation in water depth and channel cross-section, provides a poor habitat for wetland flora and fauna. Geomorphology has been applied to 'river restoration' to recreate an integrated river management strategy within artificially created river systems, maximizing biodiversity while controlling river-flow conditions.

Applied geomorphology uses a holistic approach to river response at a catchment-wide scale; the basis here is an understanding of the relationships between river form and process, sediment transport, and the important role of river-bank vegetation. Certain landscapes have specific properties that impinge on our use and development of the environment. In cold environments, the presence of ground ice leads to problems in construction, communication, and housing. In permafrost zones, the ground is permanently frozen except for the upper layers of the soil, which thaw in the summer.

The upper soil, known as the active layer, is subject to repetitive cycles of freezing and thawing, making it geomorphologically active. The ground within the active layer will suffer heaving and deformation, disrupting communications and making road construction impracticable. Applied geomorphology can be used in mapping the active layer and ground ice in areas with differing rocks and sediments.

This information is then used to evaluate the problems that are likely to affect these areas. Ground heaving depends on the depth of the active layer and the type of sediment present; fine-grained silts present more of a problem than gravels. Additional problems in permafrost areas, as, for example, in some regions of Alaska, occur where structures have suffered dramatic subsidence as a result of heating in buildings.

Without appropriate insulation, heat radiates downwards from the building into the ground, thaws the underlying ice, and increases the depth of the active layer, thus effectively changing the structure of the soil. Applied geomorphology is consequently essential in land-use planning and site evaluation, in order to recognize such potential problems as land subsidence, slope instability, invasion of windblown sand, and impacts on natural drainage systems.

Land used for agricultural production may suffer from degradation and desertification as a result of soil erosion, landsliding, and over-extraction of water for irrigation. Much agricultural practice focuses on maximizing yield and profit, often using techniques that can be detrimental to the environment, both in the short and the long term.

Applied geomorphology uses an understanding of the relationships between surface conditions, climate, vegetation, and soil erosion to advise farmers and politicians on how to improve land management for sustainable use of land and water resources. Natural hazards such as volcanic eruptions, earthquakes, and mudflows present a significant risk to the population of the surrounding area. Geomorphological mapping can be used to assess the present condition of the landscape and provide a hazard map.

The expression of a disaster may result in one settlement having significantly different risk assessment. For example, a volcanic eruption may pose a threat from volcanic ash and lava flows, pyroclastic flows, and bombardment from superheated volcanic bombs or associated hazards such as mudflows, depending on topography, soil cover, type of eruption, and predominant wind direction.

This application of geomorphological analysis is of significant interest to the emergency services and the insurance industry. Applied geomorphology can be used in modelling change to landforms and surface processes.

This can include change from human impact on the environment to future prediction of climate change, from short-term El Niño and tropical storm events to longer-term change resulting from greenhouse warming and rising sea levels. In this way, applied geomorphology has a key role in managing the environment to minimize potential degradation of land, water, and natural resources.

Geomorphology-Real-life Applications

Subsidence

Subsidence refers to the process of subsiding (settling or descending), on the part of either an air column or the solid earth, or, in the case of solid earth, to the resulting formation or depression. Subsidence in the atmosphere is discussed briefly in the entry Convection. Subsidence that occurs in the solid earth, known as geologic subsidence, is the settling or sinking by a body of rock or sediment. (The latter can be defined as material deposited at or near Earth's surface from a number of sources, most notably preexisting rock.)

As noted earlier, many geomorphologic processes can be caused either by nature or by human beings. An example of natural subsidence takes place in the aftermath of an earthquake, during which large areas of solid earth may simply drop by several feet. Another example can be observed at the top of a volcano some time after it has erupted, when it has expelled much of its material (i.e., magma) and, as a result, has collapsed.

Natural subsidence also may result from cave formation in places where underground water has worn away limestone. If the water erodes too much limestone, the ceiling of the cave will subside, usually forming a sinkhole at the surface.

The sinkhole may fill with water, making a lake; the formation of such sinkholes in many spots throughout an area (whether the sinkholes become lakes or not), is known as karst topography.

In places where the bedrock is limestone—particularly in the sedimentary basins of rivers—karst topography is likely to develop. The United States contains the most extensive karst region in the world, including the Mammoth cave system in Kentucky.

Karst topography is very pronounced in the hills of southern China, and karst landscapes have been a prominent feature of Chinese art for centuries. Other extensive karst regions can be found in southern France, Central America, Turkey, Ireland, and England.

Man-made Subsidence

Man-made subsidence often ensues from the removal of groundwater or fossil fuels, such as petroleum or coal. Groundwater removal can be perfectly safe, assuming the area experiences sufficient rainfall to replace, or recharge, the lost water. If recharging does not occur in the necessary proportions, however, the result will be the eventual collapse of the aquifer, a layer of rock that holds groundwater.

In so-called room-and-pillar coal mining, pillars, or vertical columns, of coal are left standing, while the areas around them are extracted. This method maintains the ceiling of the "room" that has been mined of its coal. After the mine is abandoned, however, the pillar eventually may experience so much stress that it breaks, leading to the collapse of the mined room. As when the ceiling of a cave collapses, the subsidence of a coal mine leaves a visible depression above ground.

Uplift

As its name implies, uplift describes a process and results opposite to those of subsidence. In uplift the surface of Earth rises, owing either to a decrease in downward force or to an increase in upward force. One of the most prominent examples of uplift is seen when plates collide, as when India careened into the southern edge of the Eurasian landmass some 55 million years ago. The result has been a string of mountain ranges, including the Himalayas, Karakoram Range, and Hindu Kush, that contain most of the world's tallest peaks.

Plates move at exceedingly slow speeds, but their mass is enormous. This means that their inertia (the tendency of a moving object to keep moving unless acted upon by an outside force) is likewise gargantuan in scale. Therefore, when plates collide, though they are moving at a rate equal to only a few inches a year, they will keep pushing into each other like two automobiles crumpling in a head-on collision. Whereas a car crash is over in a matter of seconds, however, the crumpling of continental masses takes place over hundreds of thousands of years.

When sea floor collides with sea floor, one of the plates likely will be pushed under by the other one, and, likewise, when sea floor collides with continental crust, the latter will push the sea floor under. This results in the formation of volcanic mountains, such as the Andes of South America or the Cascades of the Pacific Northwest, or volcanic islands, such as those of Japan, Indonesia, or Alaska's Aleutian chain.

Isostatic Compensation

In many other instances, collision, compression, and extension cause uplift. On the other hand, as noted, uplift may result from the removal of a weight. This occurs at the end of an ice age, when glaciers as thick as 1.9 mi. (3 km) melt, gradually removing a vast weight pressing down on the surface below.

This movement leads to what is called isostatic compensation, or isostatic rebound, as the crust pushes upward like a seat cushion rising after a person is longer sitting on it. Scandinavia is still experiencing uplift at a rate of about 0.5 in. (1 cm) per year as the after-effect of glacial melting from the last ice age. The latter ended some 10,000 years ago, but in geologic terms this is equivalent to a few minutes' time on the human scale.

Islands

Geomorphology, as noted earlier, is concerned with landforms, such as mountains and volcanoes as well as larger ones, including islands and even continents. Islands present a particularly interesting area of geomorphologic study. In general, islands have certain specific characteristics in terms of their land structure and can be analysed from the standpoint of the geosphere, but particular islands also have unique ecosystems, requiring an interdisciplinary study that draws on botany, zoology, and other subjects. In addition, there is something about an island that has always appealed to the human imagination, as evidenced by the many myths, legends, and stories about islands. Some examples include Homer's *Odyssey,* in which the hero Odysseus visits various islands in his long wanderings; Thomas More's *Utopia,* describing an idealized island republic; *Robinson Crusoe,* by Daniel Defoe, in which the eponymous hero lives for many years on an island with no companion but the trusty native Friday; *Treasure Island,* by Robert Louis Stevenson, in which the island is the focus of a treasure hunt; and Mark Twain's *Adventures of Huckleberry Finn,* depicting Jackson Island in the Mississippi River, to which Huckleberry Finn flees to escape "civilization."

One of the favourite subjects of cartoonists is that of a castaway stranded on a desert island, a mound of sand with no more than a single tree. Movies, too, have long portrayed scenarios, from the idyllic to the brutal, that take place on islands, particularly deserted ones, a notable example being *Cast Away* (2000).

A famous line by the English poet John Donne (1572-1631) warns that "no man is an island," implying that many wish they could enjoy the independence suggested by the concept of an island. Within the Earth system, however, nothing is fully independent, and, as we shall see, this is certainly the case where islands are concerned.

The Islands of Earth

Earth has literally tens of thousands of islands. Just two archipelagos (island chains), those that make up the Philippines and Indonesia, include thousands of islands each.

While there are just a few dozen notable islands on Earth, many more dot the planet's seas and oceans. The largest are these:

- Greenland (Danish, northern Atlantic): 839,999 sq. mi.(2,175,597 sq km)
- New Guinea (divided between Indonesia and Papua New Guinea, western Pacific): 316,615 sq. mi. (820,033 sq km)
- Borneo (divided between Indonesia and Malaysia, western Pacific): 286,914 sq. mi. (743,107 sq km)
- Madagascar (Malagasy Republic, western Indian Ocean): 226,657 sq. mi. (587,042 sq km)
- Baffin (Canadian, northern Atlantic): 183,810 sq. mi. (476,068 sq km)

- Sumatra (Indonesian, northeastern Indian Ocean): 182,859 sq. mi. (473,605 sq km).

The list could go on and on, but it stops at Sumatra because the next-largest island, Honshu (part of Japan), is less than half as large, at 88,925 sq. mi. (230,316 sq km). Clearly, not all islands are created equal, and though some are heavily populated or enjoy the status of independent nations (e.g., Great Britain at number eight or Cuba at number 15), they are not necessarily the largest. On the other hand, some of the largest are among the most sparsely populated.

Of the 32 largest islands in the world, more than a third are in the icy northern Atlantic and Arctic, with populations that are small or practically nonexistent.

Greenland's population, for instance, was just over 59,000 in 1998, while that of Baffin Island was about 13,200. On both islands, then, each person has about 14 frozen sq. mi. (22 sq km) to himself or herself, making.

Continents, Oceans, and Islands

Australia, of course, is not an island but a continent, a difference that is not related directly to size. If Australia *were* an island, it would be by far the largest. Australia is regarded as a continent, however, because it is one of the principal landmasses of the Indo-Australian plate, which is among a handful of major continental plates on Earth.

Whereas continents are more or less permanent (though they have experienced considerable rearrangement over the eons), islands come and go, seldom lasting more than 10 million years. Erosion or rising sea levels remove islands, while volcanic explosions can create new ones, as when an eruption off the coast of Iceland resulted in the formation of an island, Surtsey, in 1963.

Islands are of two types, continental and oceanic. Continental islands are part of continental shelves (the submerged, sloping ledges of continents) and may be formed in one of two ways.

Rising ocean waters either cover a coastal region, leaving only the tallest mountains exposed as islands or cut off part of a peninsula, which then becomes an island. Most of Earth's significant islands are continental and are easily spotted as such, because they lie at close proximity to continental landmasses.

Many other continental islands are very small, however; examples include the barrier islands that line the East Coast of the United States. Formed from mainland sand brought to the coast by rivers, these are technically not continental islands, but they more clearly fit into that category than into the grouping of oceanic islands.

Oceanic islands, of which the Hawaiian-Emperor island chain and the Aleutians off the Alaskan coast are examples, form as a result of volcanic activity on the ocean floor. In most cases, there is a region of high volcanic activity, called a hot spot, beneath the plates, which move across the hot spot. This is the situation in Hawaii, and it explains why the volcanoes on the southern islands are still active while those to the north are not: the islands themselves are moving north across the hot spot. If two plates converge and one subducts, a deep trench with a parallel chain of volcanic islands may develop. Exemplified by the Aleutians, these chains are called island arcs.

Island Ecosystems

The ecosystem, or community of all living organisms, on islands can be unique owing to their separation from continents. The number of life-forms on an island is relatively small and can encompass some unusual circumstances compared with the larger ecosystems of continents. Ireland, for instance, has no native snakes, a fact "explained" by the legend that Saint Patrick drove them away. Hawaii and Iceland are also blessedly free of serpents.

Oceanic islands, of course, tend to have more unique ecosystems than do continental islands. The number of land-based animal life-forms is necessarily small, whereas the varieties of birds, flying insects, and surrounding marine life will be greater owing to those creatures' mobility across water. Vegetation is relatively varied, given the fact that winds, water currents, and birds may carry seeds.

Nonetheless, ecosystems of islands tend to be fairly delicate and can be upset by the human introduction of new predators (e.g., dogs) or new creatures to consume plant life (e.g., sheep). These changes sometimes can have disastrous effects on the overall balance of life on islands. Overgrazing may even open up the possibility of erosion, which has the potential of bringing an end to an island's life.

AIMS AND OBJECTIVES OF GEOMORPHOLOGY

There have always been controversy and confusion about the nature of Geomorphology, these reflects the fact that historically, Geomorphologists have at one time or another attempted to answer three basic sorts of questions about the Earths landforms and landscapes (Higgins 1989), the questions are:

1. How can these features and processes be described?
2. How can they form and changed through time?
3. What processes are responsible for them and how do these processes work?

The first three lines of inquiry would seem to be a function of physical demography, one goal of which is commonly understood to be an accurate comprehensive and comprehensible description of the earth's surface. The term physiography has come to be used for such descriptive studies through the mistaken impression that the word was originally coined as a contraction of physical geography however, originally used this term in its literal sense for the study of natural phenomena in general. Later, Powell (1895) restricted it to the surface features of the earth with an emphasis on their mode of origin; and as such the term is approximately synonymous with geomorphology which has largely superseded it while W.M Davis used the term Geomorpholography for the descriptive study of landscapes. To most Geomorphologists, the ultimate aim of landform study is to explain how individual landforms and landform assemblages have originated and developed. Geomorphology was concerned primarily with the second line of inquiry i.e. the study of landform origin and change which Davis and other early authors sometimes called Geomorphology. In fact, argued that the: essential and critical distinction between Geomorphology and Dynamic geology is: the recognition of landforms or the ruminants of landforms produced by processes no longer in action, thus, in its essence and in its methodology, Geomorphology is historical.

This is the true function of the study of landforms within the generous and inclusive arms of the mother science of Geology. According to Small (1978), landforms and landscapes are so complex and pose such a variety of problems that several genetic approaches exist.

The third line of inquiry into the nature and the result of the processes that shape the earth's surface is sometimes called process Geomorphology, although the principles and methods of study are borrowed directly from the soil science and rock mechanics: hydrology and geophysics, Kirk Bryan however term such studies Dynamic Geology and emphatically excluded them from his Geomorphology.

Process Geomorphology is concerned with the investigation of the relationship process and form. This involves in the first instance a careful analysis of weathering, transport of sediments, erosion and deposition processes both as regards their mechanism and as regards rate of operation.

And secondly, relating casually, individual processes and groups of processes to particular forms. A typical example is the attempt by fluvial geomorphologists to casually relate such fluvial processes as stream discharge, bed and bank erosion, sediment transport and sediment deposition to river channel form and pattern.

Despite Bryans restriction, geomorphology is nowadays generally understood to cover all the three lines of inquiry or aspects of landform study: description, genesis and history, and process. In parts, this reflects the interdependence of all three sorts of investigations. For example, a sound reconstruction of the history of a particular landform requires both a clear picture of what that landform is today and clear understanding of the operation and results of the various processes that may have shaped it. On the other hand, a sound description of a modern landform should take into account not only its present form and structure but its antecedents as well. Finally, geomorphic studies of earth forming processes necessarily include the effects such processes have on earth materials and landforms. Such study in turn provides information needed to describe and interpret the histories of existing landscapes that may have been affected by these processes. Thus, each line of geomorphic inquiry serves the other and in turn depends on the others for fresh input observations and ideas.

10

Geomorphic Hazards : Fluvial, Coastal and Slope

Geomorphic hazards are those that originate at or near Earth's surface and include expansive soils, soil erosion, slope failures, ground subsidence and karst, river channel changes, glaciers, and coastal erosion. Geomorphic hazards are natural processes until they intersect with human activities and settlements. Geomorphic hazards can be natural or be caused/exacerbated by human activities to some degree. Expansive soils involve Vertisols with certain clay minerals or soils with a high gypsum content. They experience shrink-and-swell and deep cracking, which damages structures built on them.

Fluvial Geomorphology

Fluvial geomorphology is how rivers shape the world, and, in its most simplistic form is the interaction between sediment, water and vegetation throughout a river catchment. It is a specialist technical area which focuses on understanding and explaining river processes and how they change through time.

An understanding of river character and behaviour is key to understanding how rivers may behave in the future and under changing conditions (such as climate change or land-use changes). This information provides a holistic and long-term foundation for river management plans, flood risk assessments, erosion mitigation, gravel resource management, ecological assessments and enhancements, as well as fish passage design, stream diversion design, and stream enhancement projects.

Fluvial risk in large valleys

On European scale, the Loire (catchment: 117,000 km^2; length: 1020 km; mean discharge at Saint-Nazaire: 931 m^3/s) and the Rhône (catchment: 97,800 km^2; length: 812 km; mean discharge at Beaucaire: 1700 m^3/s) are considered large rivers. Collectively, both catchments drain 31.8% of France. They are subjected to two different types of flood, which complicates fluvial-risk management: (i) floods linked to western disturbances correspond to slow, winter flood events associated to west atmospheric circulation, leading to gradual water-table elevation and relatively long flood duration whereas (ii) Mediterranean-type floods typically correspond to flash floods generated by stormy, Mediterranean depressions occurring from the end of summer to the beginning of winter. The floods of 30 November-4 December 2003, which affected the catchments of the Rhône River with a peak discharge at Beaucaire-Tarascon of 11,500 m^3/s, and the Loire River with a peak discharge upstream to Villerest dam of 2800 m^3/s, were the latest Mediterranean-type flood events. Hydromorphological studies of these events in both catchments allow us to characterise the present fluvial hazard and risk in floodplains.

In the Loire valley, initial studies of the December 2003 flood event have been used to update the spatial limits of the flooding area. Indeed, the present atlas of the potential flood zone is based on hydrological surveys of the 19th century, but substantial morphological changes in the channel and the floodplain have since occurred. Hydrological and Differential Global Positioning System surveys associated with aerial photo interpretation have been used to define the maximum flooded area on 6 December (Grivel and Gautier, 2007). Results show the presence of unsubmerged surfaces during this flood, which was nevertheless the largest event (3400 m^3/s) in this part of the Loire valley since 1907. The hydrological impact of hydraulic structures has also been evaluated. In the middle part of the valley, the Loire River is lined on its left bank by dykes built between the 18th and the 19th centuries. These hydraulic structures were constructed, not to protect against floodwaters, but to improve conditions for navigation (Gautier *et al.*, 2007). After the large floods of the 19th century, which caused numerous dyke breaches, overflow spillways were built to reduce the hydrostatic pressure on the dykes and to facilate the gradual inundation of the floodplain, with the objective to protect urbanised zones downstream. However, most of these overflow spillways were designed for higher discharges. Therefore, the peak discharge in December 2003 propagated relatively quickly downstream and the floodwater volume was reduced only of 5-10%. Effects of river channelization by dykes have been aggravated by the recent evolution of the riverbed. In the middle part of the Loire River, active-channel width has decreased since the 19th century. In the mid-19th century, the active channel occupied 75% of the floodplain between the valley slope and the riverbank. At present, the active channel occupies only 40% of the floodplain, leading to a significant increase in the extent of islands and vegetation on the channel margins. Morphological and biological evolution has also led to a more rapid propagation of floodwaters downstream. Today, both former and current protective measures against flood risk including flood maps are used to limit or even prohibit urban development in some areas most prone to flooding.

In the Rhône valley, hydromorphological analysis of the December 2003 flood event has been used to characterise the present fluvial risk in the lower Rhône valley and to quantify fluvial processes during dyke breaching in the lower Rhône floodplain. In the alluvial plain downstream of Beaucaire-Tarascon and in the delta, the floodwaters of the Rhône River deposited a sediment volume of 810,429 m^3 (67% of which was sand) outside the dykes. The sediment balance was estimated at 674,227 m^3, taking into account eroded reaches, which represented a volume of 136,202 m^3 (Arnaud-Fassetta, 2007). This hydrological disaster was only one of six large floods that have occurred since 1993 during which peak discharges exceeded 8500 m^3/s at Beaucaire (9800 m^3/s in October 1993, 10,980 m^3/s in January 1994, 9750 m^3/s in November 1994, 8980 m^3/s in November 1996, 10 500 m^3/s in September 2002, 10,200 m^3/s in November 2002, 11,500 m^3/s in December 2003). These floods occurred in a floodplain modified by dykes and channel embankments during the second part of the 19th century. Hydraulic structures were calibrated on water levels and hydromorphological impacts of the 1856 flood (11,640 m^3/s). Active-channel contraction, incision of the channel bed and reduction of the flooding area have occurred in response to river management in the delta and the catchment, in part, in association with hydroclimatic changes following the Little Ice Age at the end of the 19th century (Arnaud-Fassetta, 2003). Land-use has significantly changed since the mid-19th century: dwellings have been built in the interfluves, in the sub-catchments and in the Rhône floodplain where flood risk is very high in case of dyke breaches. In the Rhône catchment, the concentration time of runoff and floodwaters has shortened: during the flood of December

2003 the Rhône discharge at Beaucaire increased from 2400 m^3/s to 10,000 m^3/s in ~30 hours. The maximum channel capacity of the present Rhône River, which was calibrated on the 1856 flood, was reached or even exceeded in some sections during the most recent large floods due to localised channel aggradation. Furthermore, the maximum hydrostatic pressures of the dykes are exceeded when discharge exceeds 10,000 m^3/s (HYDRATEC, 2003).In the Rhône Delta, the floodplain dykes currently protect 80,000 people, more than 2500 firms and about 70,000 hectares of agricultural lands. The financial costs of the damage caused by the flood of December 2003 (some 1.092 billion euros) demand that greater attention be given to the development of effective adaptive management strategies to reduce the impact of fluvial hazards. New forms of floodplain management such as recalibration of channel cross-sections, construction of dykes adapted to present floodwater heights and specific stream powers, harmonization of dyke systems to the sea, flood expansion areas in the lower Rhône floodplains, hydraulic restoration of Rhône palaeochannels and reduction of peak discharges upstream of Beaucaire are now encouraged more than 150 years after the first engineering works were undertaken. These questions are currently being debated between researchers, decision makers and administrators such as Regional Direction of the Environment, Rhône National Company and catchment agencies.

Slope geomorphological modelling

Computer-generated images of submarine slope landscapes can appear visually similar to their subaerial counterparts. Closer examination reveals that they have similar topographic characteristics in a quantitative sense also.

The following pages show how submarine hillsides can have similar "threshold" slopes to those found on land, how submarine canyons can have analogous variations between canyon catchment area and channel slope to those of rivers, and how smooth "diffusional" landscapes can also be found in these two different environments.

Developing "quantitative geomorphology" under water is hampered by the inaccessibility of these areas and a lack of techniques (erosion rates cannot be quantified by cosmogenic radionuclide dating, for example, because seawater rapidly absorbs cosmic radiation). This work instead explores geometrical analogies between submarine and subaerial geomorphic systems. Although the sedimentary processes causing these analogous characteristics are different, the similarity suggests that we might look for mathematical similarity in the outcomes of geomorphic processes. That in turn should guide efforts in collecting new datasets and in developing new techniques to quantify geomorphic processes.

The computer-generated image of the Southern Santa Cruz Mountains of California, an area uplifted by motion along the San Andreas fault system. Notice that the range has sharp (angular) hillcrests and the hillsides are relatively straight, not rounded. These are typical attributes of rapidly uplifting tectonic landscapes. (DEM data courtesy of the US Geological Survey.)

Threshold hillslopes

In many mountain landscapes, frequent surficial landsliding causes hillslopes to develop simple linear profiles, with gradients reflecting their limits of stability in extreme conditions of high rainfall or during seismic ground shaking. The graph on the lower right shows a series of profiles across the range divide derived from a USGS DEM (profiles are aligned with the summit of each hillslope mimicking a graph of RS Anderson (1994)). The upper graph shows a set of profiles calculated for the USA east coast continental slope (taken across inter-canyon ridges), showing a similar diversity of forms. The graph below shows averages of these profiles, revealing that, once the variability is removed by averaging, topography in both environments has linear "threshold" hillslopes. Thus the seabed morphology in these regions is sculpted by frequent surficial landslides and new accumulations of sediment are likely to be unstable, as frequently observed in core samples and from submersible.

The graphs on the right compare the relationship between channel gradient and rainfall catchment area for mountain rivers in Taiwan (based on data published by K Whipple) with similar graphs calculated for the continental slope canyons, deriving catchment area in the same way as with river drainage basins. Surprisingly the two sets of data can show a comparable relationship (graph slope) between gradient and area. The submarine data however tend to show a greater diversity of forms. In these examples (from the Virginia slope), the graphs tend to curve below A<10^7 m^2. This project aims to resolve the origin of this relationship, drawing analogies with erosion of river systems.

Rates of submarine erosion

Rates of erosion (exhumation) are difficult to ascertain because the material representing the erosion history is usually absent or difficult to relate to the erosional terrain. There are tools in subaerial geomorphology to address this issue, such as cosmogenic radionuclide dating of exposed surfaces, fission-track and He dating, other dating of strath surfaces and terraces, and sediment budgets. In the submarine landscapes, these tools are mostly not available or at least difficult to apply. This project made a first attempt using submarine slopes of volcanoes in the Canary islands, where the different ages of the volcanoes provide a chronology. We found that the greater depth of incision around Tenerife compared with El Hierro (a younger volcanic island) was compatible with the slow denudation rates seen in subaerial lowland landscapes. Rates are larger than found in the deep sea but smaller than subaerial tectonically active landscapes.

Slopes

Slope aspect is the compass bearing that a slope faces looking down slope. It is recorded either in degrees, accounting for declination, or as a general compass orientation. The direction is expressed as an angle between 0 and 360 degrees (measured clockwise from true north) or as a compass point, such as east or north-northwest.

Aspect can substantially impact local ecosystems. The impact generally increases as slope gradient and latitude increase. In the mid latitudes of the conterminous United States, this effect becomes particularly important on slopes of approximately 6 to 8 percent or greater. Increased or decreased solar radiation on slopes due to aspect can affect water dynamics across a site. In the northern hemisphere, north-northeast aspects reduce evapotranspiration and result in greater soil moisture levels, improved plant growth and biomass production, higher carbon levels, and improved drought survival rates for plants. Increased solar radiation on south-southwest aspects increases evapotranspiration and decreases biomass production, seedling survival rates, and drought survival rates for plants.

Slope gradient is the inclination of the land surface with respect to the horizontal plane. It is also commonly referred to as "slope percent" or simply "slope." It is calculated as the vertical distance divided by the horizontal distance ("rise over run"), multiplied by 100, determined at a point along a line oriented up and down slope. It directly controls the kinetic energy, erosive power, and sediment carrying capacity of running water (as overland flow or channel flow), all of which increase with increasing gradient. It inversely affects the amount of time that internal soil water is present. Many soil conservation practices, such as conservation terraces, are designed primarily to reduce slope gradient to minimize soil erosion and increase infiltration. Slope gradient also directly affects land management practices by limiting ranges of operation for various types of equipment, such as tractors and log skidders.

Slope complexity is the relative linearity or smoothness (simple) or irregularity (complex) of the ground surface leading down slope and through the point or map unit of interest. Simple slopes allow the maximum slope length with comparatively unimpeded slope wash processes. In contrast, complex slopes are composed of a series of steps commonly associated with bedrock-controlled benches or other stepped surfaces. These localized breaks in slope reduce slope length, alter slope wash processes, and commonly correspond to changes in soil types. In many places, internal soil properties are more closely related to the slope complexity than to the gradient. Slope complexity has an important influence on the amount and rate of runoff and on sedimentation associated with runoff. It can also affect soil temperature through local variation in soil aspect.

FLUVIAL LANDFORMS

Fluvial landforms are those generated by running water, mainly rivers. The term fluvial derives from the Latin word *fluvius* that means river. Fluvial landforms cover an enormous range of dimensions, from small features like rills to major continental-scale morpho-hydrological units like large rivers and their drainage basins.

The River Nile is 6,650 km long and the drainage basin of the Amazon River covers 7,050,000 km^2, an area almost as big as Australia. Rivers and streams drain most of the continental surface and occur in most environments, with the exception of some hyperarid regions, including vast sand seas, permanently frozen regions and karst terrains. Rivers flowing to the oceans drain about 68 % of the Earth's land surface. The drainage network plays a fundamental role in the continuous transfer of water and sediment from upland to lowland areas and from the continents to the oceans. Rivers transport approximately 75,000 million tonnes of material each year, of which 20,000 million tonnes reaches the sea, around 80 % in solid form and 20 % in solution. Fluvial systems are frequently the main agents involved in landscape evolution and exert a prime influence on other interrelated geomorphic systems like hillslopes, alluvial fans, deltas or beaches.

River Systems and Fluvial Landforms

Fluvial systems are dominated by rivers and streams. Stream erosion may be the most important geomporphic agent. Fluvial processes sculpt the landscape, eroding landforms, transporting sediment, and depositing it to create new landforms. Human civilization and ecosystems alike are dependent on fluvial systems.

Rivers provide water for hydroelectric power and shipping, as well as supporting stream-side wetlands (riparian areas) that are critical for clean water and provide rich habitat.

The drainage basin or watershed is a fundamental landscape unit in fluvial geomorphology. A dranage basin contains a primary, or trunk, river and its tributaries. Watersheds are separated from their neighbors by a divide; a highpoint where water flows in different directions on either side.

- Upper Basin
- Headwaters
- Mid-basin
- Low gradiant valleys and flood plains
- Lower Basin
- Depositional Zone

Floodplain Landforms

In addition to the streams themselves, the depositional habits of fluvial systems produce striking landforms. Fluvial deposits are sediments deposited by the flowing water of a stream.

A floodplain is the relatively flat surface adjacent to the river or stream. During floods, when the stream overflows its banks, water flows over the floodplain and deposits sediment. Through fluvial processes, streams construct floodplains that accommodate their maximum flood capacity. Geomorphic features of the floodplain include:

- Natural Levees—River may be immediately flanked by a buildup of sediment that forms natural levees. These provide some defense against flooding, but are occaisionally breached in areas producing flood-plain splays—coarse fan-shaped deposit of sediment created during high flow events.
- Oxbows and oxbow lakes.
- Point Bars, features of a Meandering Stream Channel.
- Terraces

Stream Channel Types

Within a single stream we can often recognize three different channel types. These unique channel types develop in response to changes in stream velocity, sediment texture, and stream grade.

Channels located in the upper reaches of many streams tend to be narrow with flow moving at high velocities. The high flow velocities found in these streams are the result of a steep grade and gravity. Within these stream systems, erosion is a very active process as the channel tries to adjust itself to the topography of the landscape. Deposition occurs primarily during periods of low flow. As a result, floodplain deposits are very limited, and the stream bed is very transient and shallow.

Figure: Upper reach of a stream in the Rocky Mountains, Canada.

Streams with high sediment loads that encounter a sudden reduction in flow velocity generally have a braided channel type. This type of stream channel often occurs further down the stream profile where the grade changes from being steep to gently sloping. In a braided stream, the main channel divides into a number of smaller, interlocking or braided channels. Braided channels tend to be wide and shallow because bedload materials are often coarse (sands and gravels) and non-cohesive. Meandering channels form where streams are flowing over a relatively flat landscape with a broad floodplain. Technically, a stream is said to be meandering when the ratio of actual channel length to the straight line distance between two points on the stream channel is greater than 1.5. Channels in these streams are characteristically U-shaped and actively migrate over the extensive floodplain.

Stream Channel Features

Within the stream channel are a variety of sedimentary beds and structures. Many of these features are dependent upon the complex interaction between stream velocity and sediment size.

Streams carrying coarse sediments develop sand and gravel bars. These types of bars seen often in braided streams which are common in elevated areas. Bars develop in braided streams because of reductions in discharge. Two conditions often cause the reduction in discharge: reduction in the gradient of the stream and/or the reduction of flow after a precipitation event or spring melting of snow and ice.

Figure: Braided stream channel with gravel bars.

Point bars develop where stream flow is locally reduced because of friction and reduced water depth. In a meandering stream, point bars tend to be common on the inside of a channel bend. In straight streams, bar-like deposits can form in response to the thalweg and helical flow.

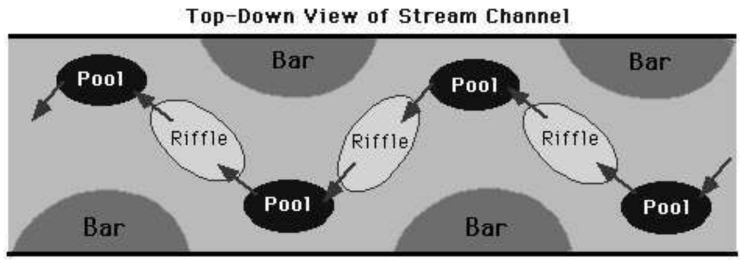

***Figure:** Overhead view of the depositional features found in a typical straight stream channel.*

In this straight channel stream, bars form in the regions of the stream away from the thalweg. Riffles, another type of coarse deposit, develop beneath the thalweg in locations where the faster flow moves vertically up in the channel. Between the riffles are scoured pools where material is excavated when the zone of maximum stream velocity approaches the stream's bed. The absolute spacing of these features varies with the size of the channel. However, the relative distance between one riffle and the next is on average five to seven times the width of the channel (exaggerated in diagram). Both of these features can also occur in sinuous channels.

Dunes and ripples are the primary sedimentary features in streams whose channel is composed mainly of sand and silt. Dunes are about 10 or more centimeters in height and are spaced a meter or more apart. They are common in streams with higher velocities. Ripples are only a few centimeters in height and spacing, and are found in slow moving streams with fine textured beds. Both of these features move over time, migrating down stream. Material on the gently sloping stoss-side of these features rolls and jumps up the slope under the influence of water flow. Particles move up the slope until they reach the crest of the feature and then avalanche down the steeper lee-side to collect at the base of the next dune or ripple. This process is then repeated over and over again until the material reaches a location down stream where it is more permanently deposited.

The Floodplain

Alongside stream channels are relatively flat areas known as floodplains. Floodplains develop when streams over-top their levees spreading discharge and suspended sediments over the land surface during floods.

Figure: The following Landsat 5 image taken in September 1992 shows a section of the Missouri River at Rocheport, Missouri. The oblique perspective of this image is looking westward or upstream. This image has been color enhanced and modified to show an exaggerated topographic relief. Bare soil and plowed land appears red, vegetation appears green, and water is dark blue. A flat river flood plain can be seen in the center of the image. Because of the season, most of the farmland located on the rich and fertile soils of the floodplain is plowed and devoid of vegetation.

Levees are ridges found along the sides of the stream channel composed of sand or gravel. Levees are approximately one half to four times the channel width in diameter. Upon retreat of the flood waters, stream velocities are reduced causing the deposition of alluvium.

Repeated flood cycles over time can result in the deposition of many successive layers of alluvial material. Floodplain deposits can raise the elevation of the stream bed. This process is called aggradation.

Floodplains can also contain sediments deposited from the lateral migration of the river channel. This process is common in both braided and meandering channels. Braided channels produce horizontal deposits of sand during times of reduced discharge. In meandering streams, channel migration leads to the vertical deposition of point bar deposits. Both braided and meandering channel deposits are more coarse than the materials laid down by flooding.

A number of other geomorphic features can be found on the floodplain. Intersecting the levees are narrow gaps called crevasses. These features allow for the movement of water to the floodplain and back during floods. Topographical depressions are found scattered about the floodplain. Depressions contain the some of the finest deposits on the floodplain because of their elevation. Oxbow lakes are the abandoned channels created when meanders are cut off from the rest of the channel because of lateral stream erosion.

Alluvial Fans and Deltas

Streams flowing into standing water normally create a delta. A delta is body of sediment that contains numerous horizontal and vertical layers. Deltas are created when the sediment load carried by a stream is deposited because of a sudden reduction in stream velocity. The surface of most deltas is marked by small shifting channels that carry water and sediments away from the main river channel. These small channels also act to distribute the stream's sediment load over the surface of the delta. Some deltas, like the Nile, have a triangular shape. Streams, like the Mississippi, that have a high sediment content and empty into relatively calm waters cause the formation of a birdfoot shaped delta.

Figure: Nile Delta (Source: NASA).

Most deltas contain three different types of deposits: foreset, topset and bottomset beds. Foreset beds make up the main body of deltas.

They are deposited at the outer edge of the delta at an angle of 5 to 25 degrees. Steeper angles develop in finer sediments. On top of the foreset beds are the nearly horizontal topset beds.

These beds are of varying grain sizes and are formed from deposits of the small shifting channels found on the delta surface. In front and beneath the foreset beds are the bottomset beds. These beds are composed of fine silt and clay. Bottom set beds are formed when the finest material is carried out to sea by stream flow.

An alluvial fan is a large fan-shaped deposit of sediment on which a braided stream flows over (10z-10). Alluvial fans develop when streams carrying a heavy load reduce their velocity as they emerge from mountainous terrain to a nearly horizontal plain. The fan is created as braided streams shift across the surface of this feature depositing sediment and adjusting their course. The image below shows several alluvial fans that formed because of a sudden change in elevation.

Figure: Alluvial Fans - Brodeur Peninsula, Baffin Island, Canada.

CAUSES OF COASTAL HAZARDS

The population that lives along or near our coastlines are an extremely vulnerable population. There are numerous issues facing our coastlines and there are two main categories that these hazards can be placed under, National disasters and Human disasters. Both of these issues cause great damage to our coastlines and discussion is still ongoing regarding what standards or responses need to be met to help both the individuals who want to continue living along the coastline, while keeping them safe and not eroding more coastline away. Natural disasters are disasters that are out of human control and are usually caused by the weather.

Disasters that include but are not limited to; storms, tsunamis, typhoons, flooding, tides, waterspouts, nor'easters, and storm surge. Human disasters occur when humans are the main culprit behind why the disaster happened. Some human disasters are but are not limited to; pollution, trawling, and human development. Natural and human disasters continue to harm the coastlines severely and they need to be researched in order to prepare/stop the hazards if possible.

The populations that live near or along the coast experience many hazards and it affects millions of people. Around ten million people globally feel the effects of coastal problems yearly and most are due to certain natural hazards like coastal flooding with storm surges and typhoons. A major problem related to coastal regions deals with how the entire global environment is changing and in response, the coastal regions are easily effected.

Storms, Flooding, Erosion

Storms are one of the major hazards that are associated to coastal regions. Storms, flooding, and erosion are closely associated and can happen simultaneously. Tropical storms or Hurricanes especially can devastate coastal regions. For example, Florida during Hurricane Andrew occurred in 1992 that caused extreme damage.

It was a category five hurricane that caused $26.5 billion in damages and even 23 individuals lost their lives from the storm. Hurricane Katrina also caused havoc along the coast to show the extreme force a hurricane can do in a certain region. In almost all cases, storms are the major culprit that causes flooding and erosion. Flash flooding is caused by storms that occurs when a massive amount of rainfall comes down into an area over a short period of time. Where as a storm surge, which is closely related to tropical storms, is when the wind collects and pushes water towards low pressure or inland and can rise rapidly. It is an offshore rise of water and overall creates a higher sea level that rises and is pushed inland. The amount of rise or fall of storm surge depends greatly on the amount and duration of wind and water in a specific location. Also if it occurs during a high tide it can have an even greater effect on the coast.

Almost all storms with high wind and water cause erosion along the coast. Erosion occurs when but not limited to; along shore currents, tides, sea level rise and fall, and high winds. Larger amounts of erosion cause the coastline to erode away at a faster rate and can leave people homeless and leave less land to develop or keep for environmental reasons. Coastal erosion has been increasing over the past few years and it is still on the rise which makes it a major coastline hazard. In the United States, 45 percent of its coast line is along the Atlantic or Gulf coast and the erosion rate per year along the Gulf coast is at six feet a year. The average rate of erosion along the Atlantic is around two to three feet a year. Even with these findings, erosion rates in specific locations vary because of various environmental factors such as major storms that can cause major erosion upwards to 100 feet or more in only one day.

Pollution, Trawling, Human Development

Pollution, trawling, and human development are major human disasters that effect coastal regions. There are two main categories related to pollution, point source pollution, and nonpoint source pollution. Point source pollution is when there is an exact location such as a pipeline or a body of water that leads into the rivers and oceans. Known dumping into the ocean is also another point source of pollution. Nonpoint source pollution would pertain more to fertilizer runoff, and industrial waste. Examples of pollution that effect the coastal regions are but are not limited to; fertilizer runoff, oil spills, and dumping of hazardous materials into the oceans. More human acts that hurt the coastline are as follows; waste discharge, fishing, dredging, mining, and drilling. Oil spills are one of the most hazardous dangers towards coastal communities. They are hard to contain, difficult to clean up, and devastate everything. The fish, animals such as birds, the water, and especially the coastline near the spill. The most recent oil spill that had everybody concerned with oil spill was the BP oil spill.

Trawling hurts the normal ecosystems in the water around the coastline. It depletes all ecosystems on the ocean floor such as, flounder, shellfish, marsh etc.. It is simply a giant net that is drug across the ocean floor and destroys and catches anything in its path. Human development is one of the major problems when facing coastal hazards. The overall construction of buildings and houses on the coast line takes away the natural occurrences to handle the fluctuation in water and sea level rise. Building houses in pre-flood areas or high risk areas that are extremely vulnerable to flooding are major concerns towards human development in coastal regions. Having houses and buildings in areas that are known to have powerful storms that will create people to be in risk by living there. Also pertaining to barrier islands, where land is at risk for erosion but they still continue to build there anyway. More and more houses today are being taken by the ocean, look at picture above.

HILL SLOPES GLACIAL

High Asian glacial landscapes have large variations in topographic relief and the size and steepness of snow accumulation areas. Associated differences in glacial cover and dynamics allow a firstorder determination of the dominant processes shaping these landscapes.

Here we provide a regional synthesis of the topography and flow characteristics of 287 glaciers across High Asia using digital elevation analysis and remotely sensed glacier surface velocities. Glaciers situated in lowrelief areas on the Tibetan Plateau are mainly nourished by direct snowfall, have little or no debris cover, and have a relatively symmetrical distribution of velocities along their length.

In contrast, avalanchefed glaciers with steep accumulation areas, which occur at the deeply incised edges of the Tibetan Plateau, are heavily covered with supraglacial debris, and flow velocities are highest along short segments near their headwalls but greatly reduced along their debrismantled lower parts. The downstream distribution of flow velocities suggests that the glacial erosion potential is progressively shifted upstream as accumulation areas get steeper and hillslope debris fluxes increase.

The hillslope legacies of glacial systems are moraines and the debris on and in the glaciers. Material that is eroded from glacial headwalls in the accumulation zone ultimately lands on the glaciers below, where it forms supraglacial debris once it is exposed in the ablation zone [e.g., Boulton and Eyles, 1979; Hewitt, 2009]. At debris thicknesses of <<“2 cm, the lower albedo of debris compared to clean ice absorbs more incoming radiation energy and thus increases melt rates. In contrast, the insulating effect of debris cover dominates at thicknesses ><“2 cm and melt rates decrease to values below those of clean ice [e.g., Østrem, 1959; Mattson et al., 1993; Kayastha et al., 2000]. Because debris cover strongly modulates melt rates [e.g., Ogilvie, 1904; Østrem, 1959] it also influences glacial mass balances and thus the size and geometry of glaciers. Based on these arguments, debris cover should therefore also have an effect on the magnitude and distribution of glacial velocities, which are firstorder factors controlling their erosion potential.

Glacial lakes and ponds

Lakes and ponds may also be caused by glacial movement. Kettle lakes form when a retreating glacier leaves behind an underground or surface chunk of ice that later melts to form a depression containing water. Moraine-dammed lakes occur when glacial debris dam a stream (or snow runoff). Jackson Lake and Jenny Lake in Grand Teton National Park are examples of moraine-dammed lakes, though Jackson Lake is enhanced by a man-made dam.

- Kettle lake: Depression, formed by a block of ice separated from the main glacier, in which the lake forms
- Tarn: A lake formed in a cirque by overdeepening
- Paternoster lake: A series of lakes in a glacial valley, formed when a stream is dammed by successive recessional moraines left by an advancing or retreating glacier
- Glacial lake: A lake that formed between the front of a glacier and the last recessional moraine

Ice features

Apart from the landforms left behind by glaciers, glaciers themselves may be striking features of the terrain, particularly in the polar regions of Earth. Notable examples include valley glaciers where glacial flow is restricted by the valley walls, crevasses in the upper section of glacial ice, and icefalls—the ice equivalent of waterfalls.

Disputed origin

The glacial origin of some landforms has been questioned.

Erling Lindström has advanced the thesis that roches moutonnées may not be entirely glacial landforms taking most of their shape before glaciation. Jointing that contribute to the shape typically predate glaciation and roche moutonnée-like forms can be found in tropical areas such as East Africa and Australia. Further at Ivö Lake in Sweden weathered rock surfaces exposed by kaolin mining resemble roche moutonnée.

The idea of elevated flat surfaces being shaped by glaciation—the glacial buzzsaw effect—has been rejected by various scholars. In the case of Norway the elevated paleic surface has been proposed to have been shaped by the glacial buzzsaw effect. However, this proposal is difficult to reconcile with the fact that the paleic surface consist of a series of steps at different levels. Further glacial cirques, that in the buzzsaw hypothesis contribute to belevel the landscape, are not associated to any paleosurface levels of the composite paleic surface, nor does the modern equilibrium line altitude (ELA) or the Last Glacial Maximum ELA match any given level of the paleic surface. The elevated plains of West Greenland are also unrelated to any glacial buzzsaw effect.

The Gulf of Bothnia and Hudson Bay, two large depressions at the centre of former ice sheets, are known to be more the result of tectonics than of any weak glacial erosion.

11

Applied Fluvial Geomorphology

Fluvial Geomorphology

An understanding of river- and stream-channel geomorphic responses to various human-caused and natural disturbances is important for effective management, conservation, and rehabilitation of rivers and streams to accommodate multiple, often conflicting, needs. Channel changes may have implications for the protection of property and structures, water supply, navigation, and habitat.

The channel-bank erosion that accompanies natural channel migration on a flood plain represents a constant threat to property and structures located in or near the channel. Various human-caused and natural disturbances introduce additional instability to which rivers and streams adjust. Human-caused disturbances include reservoirs, channelization, in-channel sand and gravel extraction, and urbanization. A common natural disturbance is a flood. Possible geomorphic responses of a channel to disturbances include channel-bed degradation (erosion), channel-bed aggradation (deposition of material), channel widening, and channel straightening. These adjustments represent the channel's attempt to establish a new approximate equilibrium condition.

Erosion along Stranger Creek has reached a house along the bank

Channel adjustments are a concern for several reasons. A substantial lowering of the channel bed poses an immediate threat to bridge pier foundations as well as buried pipelines and cables. In addition, substantial bed lowering increases bank height and bank instability that may trigger channel widening. Channel aggradation raises the bed elevation, reduces channel capacity, and increases the likelihood of flooding. Any channel-bed changes that occur on the main-stem rivers and streams also may migrate upstream on the tributaries where additional property, structures, and habitat may be at risk. Finally, any long-term channel adjustment processes also may instigate or worsen local scour problems.

Geomorphic investigations conducted by the U.S. Geological Survey since 1995 have mostly focused on the response of river and stream channels to various types of natural and human-caused disturbances including floods, reservoir construction and operation, and channelization. Such studies document channel changes, reconstruct historical conditions, determine the causes of channel changes, estimate the rate of geomorphic processes, and, in some cases, can enable predictions of future channel changes.

Methods have included the use of streamgage data, multidate aerial photography, and onsite data collection to determine the location, timing, magnitude, direction, duration, and rate of channel change.

FLUVIAL PROCESSES AND LANDFORMS

Geomorphic significance of fluvial systems tied to the hydrological cycle

- water evaporated from the ocean surface transported over continental regions where it condenses and precipitates
- gravity returns some of this water back to the oceans as runoff
- streams remaove excess water from the continents
- streams erode, transport and eventually deposit weathered materials to form fluvial landforms
- the geomorphic effect of flowing water displays significant spacial variations in its nature and intensity related to climatic and vegetative patterns
- fluvial geomorphic features are also related to local soil and bedrock characteristics
- flowing water is most effective as a landcape modifier in semiarid and subhumid regions
- humid regions with abundant precipitation have only a moderate rate of development of fluvial features due to the protective influence of vegetation
- the climatic zone of maximum fluvial effectiveness is one that receives enough rainfall to produce considerable storm runoff, but dry enough to support only a limited vegetation cover.

Channel Morphology

The shape that a stream takes, or its channel morphology, is a function of the sediment carried and deposited by the stream. This divides medium to low gradiant streams into two general categories, meandering and braided.

Causes and Characteristics of Stream Flow

- land surfaces have infiltration capacities (maximum rate at which they can absorb water)
- water not absorded immediately collects at the surface (filling in depressions)
- if the land surface is sloped then water will begin to flow downhill under the influence of gravity
- flows increase in volume as the total upslope area being drained becomes progressively larger
- Channelization occurs where the water depth and speed enable flow to overcome the cohesion of soil particles
- on unprotected slopes (cleared of vegetation for agriculture) a parallel series of channels (rills) form first
- when rills coalesce they form steep sided gullies
- when gullies deepen enough to reach the water table of the area they begin tapping into groundwater supplies at which point they become a perennial stream channel
- rills and gullies normally do not form on undisturbed, well vegetated slopes

Drainage Systems and Patterns

- streams are organized into drainage systems (numerous interconnected stream segments)
- these stream segments act collectively to remove excess water and sediments from a drainage basin (total area drained by the stream system)
- drainage basin: main river (stream) fed by smaller tributaries (analogous to tree branches)
- drainage divides: separate the flow and direct it toward adjacent drainage systems (Continental Divide)

Stream Orders

- the size of a stream described as: length, width, volume of water discharged and drainage areas

- collectively these parameters can be conveyed by stream order
- streams are ranked from 1st to 12th order
- 1st order stream isa small stream with no tributaries
- 2nd order stream is formed when two 1st order streams form
- subsequent orders are formed when when two strams of the previous order join
- the physical dimensions of the streams increase with increasing stream order
- Mississippi River is a 10th order stream,
- Amazon River is a 12th order stream

Drainage Patterns

- the surface pattern collectively formed by the streams in a drainage system
- all drainage systems in a given area generally exhibit the same general pattern
- drainage patterns respond to bedrock types and preexisting surface features
- dendritic pattern: most common drainage pattern
- develops on gently sloping surfaces that are homogeneous in their resistance to erosion
- so structural control on stream location
- trellis pattern: typical of regions with lots of ridges and valleys where resistant and nonresistant sedimentary rock types alternate
- major streams in valleys flow parallel to ridges while the smaller tributaries flow down steep ridge flanks (perpendicular to the major streams) to join the major streams (Appalachian Mountains, Valley and Ridge Regions)
- rectangular pattern: produced when the drainage pattern is controlled by intersecting fault and fracture systems (areas of crystalline rocks; granites)

- right angle bends in the streams
- radial pattern: outward flow of streams in all directions from a central pek or upland
- well developed near isolated volcanic peaks
- deranged pattern: large degree of spatial disorganization
- streams are highly irregular in their directions of flow
- they pass through depressions occupied by lakes and swamps
- areas of recent geologic orogin where drainage patterns have not had time to develop (glacially scoured regions: Canadian Shield and Scandinavia)

Erosion, Transportation and Deposition

- stream erosion results from the friction and shear forces generated between flowing water and nonmoving material with which it comes into contact
- if the shearing forces produced by the water exceed the cohesive strength of the nonmoving material, the material will be eroded and carried downstream
- 95 % of gravitational energy that produces stream flow is converted to heat through friction between water molecules
- the remaining 5 % is used for sediment erosion and transport
- erosive energy of a stream is related to its velocity, volume and the amount of friction between its banks and bed
- estimated that a stream's erosive power increases as the cube of its increase in velocity (if the speed of the flow doubles, its erosive power increases by 8 times)
- critical erosion velocity: velocity of water needed to produce enough friction to erode material and produce eddies (turbulence) within the flow to lift the material and erode them
- sediment erosion by streams is accomplished in three different ways:

- hydraulic action: direct sweeping away of loose material by friction and turbulence by moving water
- abrasion: scraping of particles carried in the water against secondary materials
- corrosion: dissolving of rock materials so that they enter the flow in a molecular state

Transportation

- turbulence lifts the sediments and keeps them from settling, allows the current to transport them downstream
- transport methods:
- dissolved load: transporting material in a dissolved state (particularly important in areas where the flow velocity is low, low turbulence)
- suspended load: consist of fine textured particles that are carried within the flow
- particles are kept from settling by the upward component of turbulent eddies generated by friction
- common where streams flow through unconsolidated areas
- bed load: the coarsest particles a stream is capable of moving
- particles are moved along the stream bed because they are too heavy for turbulence to transport them for great distances
- travel by saltation (roll slide or bounce along a stream bed)
- many bed loads are only transported during floods

Deposition

- streams deposit sediments when their velocities slow sufficiently to reduce their turbulence
- this makes them incapable of transporting all of their sediment load
- coarsest materials are deposited first followed by finer sediments
- sediments are sorted or separated by size

- accumulation of thick deposits of sediments is common in the lower portions of stream systems (deltas)
- deposition also occurs locally and temporarily at numerous points along the courses of all streams
- accumulation of fine-textured stream-deposited sediments are called alluvium

Fluvial Landforms

- weathering, mass wasting and fluvial erosion act to erode the surface of the earth and produce most gradational landforms
- weathering acts first producing unconsolidated surface materials, mass movements transport these materials to stream valleys, and streams transport these materials to the sea
- during stream transport a variety of fluvial erosional and fluvial depositional features are formed
- the dominant fluvial feature is a valley (a linear lowland produced by erosion)
- the dimensions of a valley are determined by the volume of the river flow, the rate of stream erosion and the length of time over which the valley has developed
- Erosional Features:
- erosional valley refers to a valley still being actively deepened by a river downcutting toward base level (the elevation at any point along the stream's course where the gradient is sufficient to transport its sediment load past that point where there is no erosion or deposition)(flow equilibrium)
- a stream downcutting rapidly through rock material strong enough to maintain steep slopes produces a gorge
- a wider gorge is called a canyon
- the sides of gorges and canyons are determined by the material comprising the sides
- if the rock type is homogeneous a smooth slope will develop

- if the layers of rock alternate in erosive resistance an irregular slope will develop
- erosional valleys exist mostly in upland areas (plateaus, mountainous regions) in arid or semiarid regions

Erosional Hills and Mountains

- erosional hills and mountains are the highland features that remain following the fluvial dissection of a plateau (carved out due to fluvial erosion)
- the pattern of hills or mountains within a region is determined by the fluvial drainage pattern that develops
- Appalachian and Ozark Plateaus form disorganized upland surfaces from dendritic drainage patterns
- in areas where folding and faulting have exposed rock sequences with different resistances to erosion, streams develop linear drainage patterns (trellis pattern)
- this will usually produce a parallel system of resistant linear ridgesand valleys (Appalachian Valley and Ridges)

Fluvial Depositional Features

- deposition occurs along the lower portions of drainage systems (gentle slopes)
- sediment supplies are in large volume
- stable long-term base level is more likely to be achieved than further upstream
- Floodplain is the dominant fluvial depositional feature
- floodplains contain secondary features such as meander deposits, natural levees, deltas and alluvial terraces
- floodplains are broad, flat-floored valleys covered by alluvium
- they are subject to flooding at all times of high water
- large quantities of sediment are transported by the stream during flooding and are deposited in a layer over the surface
- this adds valuable topsoil to the floodplain

Meanders

- the development of broad looping bends (meanders) in the courses of streams occupying flood planes is responsible for many floodplain features
- a meander begins as a slight bend in the channel
- the diversion in flow results in the erosion of unconsolidated material on the outside of the bend allowing it to become more pronounced
- as the bend increases in amplitude the river flow is directed toward the outer bank where erosion continues producing an undercut bank
- the meander slowly migrates down the valley as it grows
- the reduction of speed on the inner side of the developing meander causes the deposition of sediments as a point bar deposit
- continued erosion on the outer curve of the meander and deposition on the inner curve of the meander increases the amplitude of the meander
- the meander eventually forms a narrowing meander neck
- eventually meander sides intersect to form a meander cutoff
- the meander is eventually abandoned by the river (river takes the shorter route through the cutoff) and the meander is abandoned by the river to form an oxbow lake
- oxbow lakes eventually fill with sediment becoming first a wetland (swamp) then dry land

Natural Levees

- water spilling over the banks of a river during a flood is subject to reduced current speeds and increased friction
- the coarsest sediments sometimes deposit atop the river banks where they accumulate to form a natural levee
- levees are broad ridges of fine sand and coarse silt that parallel each side of the stream channel they increase the channel's stability

Deltas

- deltas are deposits of alluvium formed when streams enter standing water bodies
- deposition takes place because the sudden reduction of flow that occurs as the stream enters the standing water body causes it to lose its sediment-carrying capacity
- all but dissolved sediments and clay sized particles settle out of the river flow
- the outward flow of sediments produces a fan shaped delta
- deltas are proportional in size to the sediment loads that a river carries
- river deltas are seaward extensions of floodplains

Terraces

- a river that produces a flood plane has reached a stable base level
- if the base level is lowered again the river will continue down cutting its channel into the floodplain
- the remnant older floodplain forms a terrace with steep slopes well back from the current river channel
- a pair of terraces will be present on both sides of the river valley
- separation of the terraces on either side of the river channel depends on the width of the new river floodplain

Alluvial Fans

- these are fan-shaped deposits of alluvium that form in arid regions (deserts) when ephemeral streams (those that actually flow only rarely) flood and empty onto the desert plains out of the mountains

Glacial Landforms

Glacier Formation and Distribution

- a glacier is a mass of freshwater ice formed on land that is in motion or has been in motion in the past

- glaciers move due to gravity
- source of glacial ice is compacted snow
- glaciers develop when the accumulation of snow during the year (winter) exceeds the loss of snow by melting and sublimation (summer)
- melting and sublimation processes are referred to as ablation
- need cold temperatures (high latitudes, alpine regions) and significant snowfalls
- glacier = snowfall accumulation > ablation
- snow gradually converted to glacial ice due to burial and compaction (change in density from fresh snow to a dense granular state called firn then to glacial ice)

Global Extent of Glaciers

- glaciers cover approximately 10% of the earth's land surface
- continental glaciers - form in high latitudes and cover extensive land areas
- alpine glaciers - form in high elevations in alpine areas
- two biggest continental glaciers: Antarctica (12, 588,000 km2) and Greenland (1,802,600 km2)
- these two continental glaciers contain ~ 75% of the world's freshwater supply
- if they melted sea levels would raise 45 meters
- alpine glaciers constitute a combined area of 508,000 km2
- most significant alpine glaciers: Alaska, Canadian Rockies, Canadian Arctic Islands, Andes, European Alps, Himalayas.......

Glacier Processes and Features

- glacier motion is caused by gravity (compaction deforms ice which may melt a flow downhill)(river of ice)
- absolute motion is the actual movement of particles of ice in the glacier (very slow, usually imperceptible)

- surges are sudden increases in the rate of advance of a glacier
- surges are likely a result of abnormally high snowfall or are sometimes initiated by accumulations of subglacial meltwater
- relative motion of a glacier refers to advance or retreat of an ice front and is determined by the relationship between the absolute motion of a glacier and its rate of ablation
- a glacier that gains mass and advances usually has a positive mass balance (imputs exceed outputs)
- a glacier that loses mass and retreats usually has a negative mass balance (outputs exceed inputs)
- every glacier has a zone of accumulation and a zone of ablation (division between these zones is known as an equilibrium line)
- glacier movement has two main components: basal sliding (sliding of ice on bedrock) and internal flow (laminar flow of horizontally oriented ice crystals shearing over those below; plastic flow due to the weight of the overlying ice)
- the absolute motion of a glacier is the sum of basal sliding motion + internal flow motion
- internal flow rates are greatest near the center of the glacier than near the margins (next to bedrock)(similar to flow in a river)

Glacier Erosion and Related Features

- glaciers are very effective solid state eroders and transporters of regolith (weathered rock material) due to the tremendous pressure exerted by hundreds and even thousands of meters of ice
- Scraping is accomplished primarily by rocks frozen into the base of the ice and dragged over bedrock surfaces (large frozen in rocks sometimes produce striations or scratches in the bedrock which are valuable indicators of past ice flow direction; sand and silt sized particles smooth and polish some rock surfaces- like sandpaper)

- Plucking involves the excavation of angular, sometimes very large rock fragments from bedrock
- ice melting on the up-glacier side of an obstacle (hills, boulders) flows around or over the obstruction, freezes in rock fractures on the down-glacier side and plucks out rock fragments as the glacier moves on
- glaciers are most effective at eroding unconsolidated material (material from previous glaciations or soils) which may strip the unconsolidated cover down to bedrock
- glaciers pick up sediment, incorporate it into the ice flow and transport it away where it is eventually deposited as various features

Glacier Deposition and Related Features

- glaciers transport debris in a conveyor belt type fashion
- glacial drift is the general term for all deposits of glacial origin
- drift deposited directly by the glacier is known as till (unsorted, angular with particles ranging from clay to boulder sizes)
- glacial-fluvial sediments results from drift being reworked by glacial meltwaters after their initial deposition by glaciers (partly stratified and sorted)
- till is primarily deposited by a process of dumping at the glacier terminus (snout) or by leaving material on the surface beneath the glacier (produce deposits known as moraines)
- a terminal moraine is formed where the glacier terminates (conveying sediment to the end where it is dumped) (indicate the furthest extent of the glacier)
- when a glacier retreats, a series of moraines can remain marking the retreat of the glacier snout- these are called recessional moraines
- between terminal and recessional moraines a thin deposit of till is deposited by the retreating glacier - called a ground moraine

- drumlins are low elongated hills deposited by ice sheets (long axis of the drumlin parallel to the direction of ice movement)
- probably a result of reworking of previously deposited till by the readvancement of glacial ice
- composed of unsorted till
- usually occur in groups all parallel to one another (sometimes hundreds)
- as a glacial ice front recesses, sometimes large blocks of ice break off the main ice sheet and produce depressions called kettles in the ground moraine till
- if water from the isolated melting ice block is retained in the kettle depression, this is then known as a kettle lake
- glacio-fluvial deposits, generally called outwash deposits are usually deposited at or beyond the margins of the glacier or beneath the glacier (subglacial deposits)
- glacial meltwater is loaded with sediment that very often is deposited in braided stream deposits (multiple stream channels) across the area beyond the glacial front
- a lengthy deposit of glacio-fluvial alluvium confined to a valley is called a valley train
- sinuous ridges of glacio-fluvial deposited stratified drift are called eskers which are formed subglacially as meltwater flowing through tunnels in the interior of the ice sheet become choked off during a time when the ice sheet is stagnant (neither advancing or retreating)
- as water flow slows in these subglacial tunnels, sediment is deposited
- eskers are exposed when the glacier melts away
- kames are small, steep mounds or conical hills of stratified drift that form in pockets (rather than channels) in stagnant glacial ice through fluvial action as subglacial deltas or fans

Alpine Glaciers and Related Landforms

- most alpine glaciers originate in relatively small mountainside hollows called cirques
- if conditions are marginal for ice accumulation, the glacier may remain as a small, rounded cirque glacier
- if conditions are favourable for ice accumulation, the glacier will outgrow its cirque and flow down mountain side valleys as a long narrow river of ice called a valley glacier
- lateral moraines are stripes of rock debris that accumulate along the sides of alpine glaciers as a result of both glacial erosion and weathering along the sides of glacial valleys
- when two tongues of glacial ice meet their innermost lateral moraines coalesce to form a single medial moraine within the glacier

Alpine Glacier Erosional Landforms

- erosion by alpine glaciers tends to increase the relief and ruggedness of the topography and create a number of landform features
- cirques are bowl-shaped depressions eroded into the rock faces of mountainsides in glaciate areas
- consist of a steep headwall, a smooth central basin and a raised lip of threshold
- cirques that formed due to previous glaciations that are now devoid of ice and contain lakes are called tarns
- as cirques erode from opposite sides of a linear mountain they can eventually produce a narrow serrated ridge called an arte or erosion by three or more cirques can produce a sharp-crested pyramidal peak called a horn (Matterhorn, Switzerland)
- a glacial trough is a deep U-shaped valley, usually of fluvial origin, that has been subsequently occupied and modified by a valley glacier
- tributary glaciers entering a valley glacier (from the side) also carve out U-shaped valleys, but usually not as deep

- the result (when the ice finally melts) is intersecting hanging valleys with steep walls

- at higher latitudes, glacial troughs may be eroded below sea level and occupied by ocean water to become fiords when the ice melts

Alpine Glacier Depositional Landforms

- same features moraine and glacio-fluvial features as ice sheet landforms but less extensive and not defined as well

The Pleistocene Epoch

The Pleistocene was a time when large ice sheets advanced and retreated over North America

Ice Age Climate

Oxygen isotope records from marine core sediments suggest that global climate has gradually cooled from the beginning of the Eocene through the Pleistocene. Oxygen isotope data from deep sea cores reveal that during the past 2 million years there were at least 20 major glaciation periods.

Stratigraphic evidence indicates that there were at least 4 major episodes of Pleistocene glaciation in North America as evidenced by glacial till deposits (separated by 4 interglacials as evidenced by soil horizons).

The most recent glacial event is called the Wisconsinan Glaciation which began about 70,000 years ago.

Ice sheets that grew as a result of the cool temperatures imposed by the Wisconsinan Glaciation (Laurentide Ice Sheet in North America and the Fennoscandian Ice Sheet in Northern Europe) composed a volume of 43 million cubic kilometres during the Glacial Maximum.

Ocean Sea level was lower by approximately 130 meters.

The Laurentide Ice Sheet was composed of 4 main ice loges centred over Greenland, Baffin Island, the Keewatin region of the North West Territories and Labrador.

There was also a separate Cordilleran Ice Sheet (ice free corridor between the Cordilleran and Keewatin Ice Sheets).

During the Wisconsinan Glaciation there were intervals of ice sheet advance and retreats

Tilt cycle: 41,000 years

Wobble cycle: 23,000 years

These effects occasionally reinforce one another to cause unusually great or small net solar radiation on the earth's surface (Milankovitch cycles).

Pluvial and Proglacial Lakes

The Laurentide Ice Sheet extended the Arctic front or jet stream (the boundary between cold air from the north and warm air from the south) south over the southwestern US during the Wisconsinan Glaciation

Weather systems track along the jet stream and therefore the result of the jet stream being further south during the Wisconsinan glaciation is more precipitation in the southwestern US.

Water as a result of the increased precipitation collects in the basins defined by the Basin and Range geological province resulting in Pluvial Lakes (Lake Bonneville, Lake Lahontan, Lake Manly)

Once the Laurentide Ice sheet retreated north at the end of the Wisconsinan Glacial Period the jet stream assumed a more northerly route which reduced precipitation in southwestern US causing the Pluvial Lakes to evaporate leaving dry lake deposits in the basins (Playas)

Proglacial Lakes are formed by the meltwater accumulating along the margins of glaciers.

Numerous proglacial lakes existed during the Pleistocene (Lake Missoula, Lake Columbia, Lake Agassiz, Great Lakes)

The Scablands of Eastern Washington

Lake Missoula formed when an advancing glacier plugged the Clark Ford Valley at Ice Cork, Idaho causing glacial melt water to fill the western valleys of Montana.

When the ice dam impounding Lake Missoula failed (ice retreated back) the water drained south at a tremendous velocity draining south and southwest across Idaho and into Washington

The water carved out huge valleys in the bedrock (plucking off pieces of basalt 10 m across, channels contain giant ripple marks - 10 m high and 70 to 100 meters apart)

Lake Missoula formed, drained (producing scablands) and re-formed at least 4 times (possibly more)

Glaciers and Isostasy

When ice accumulates in great quantities in a particular area the earth's crust responds by gradually subsiding down into the mantle under the weight of the load (in some places 300 meters below previous elevations)

When the ice melts the weight of the ice is removed and the earth's crust slowly rebounds (isostatic rebounding)

Fluvial processes

In geography and geology, fluvial processes are associated with rivers and streams and the deposits and landforms created by them. When the stream or rivers are associated with glaciers, ice sheets, or ice caps, the term glaciofluvial or fluvioglacial is used.

Fluvial processes

Fluvial processes include the motion of sediment and erosion or deposition on the river bed.

Erosion by moving water can happen in two ways. Firstly, the movement of water across the stream bed exerts a shear stress directly onto the bed. If the cohesive strength of the substrate is lower than the shear exerted, or the bed is composed of loose sediment which can be mobilized by such stresses, then the bed will be lowered purely by clearwater flow. However, if the river carries significant quantities of sediment, this material can act as tools to enhance wear of the bed (abrasion). At the same time the fragments themselves are ground down, becoming smaller and more rounded (attrition). Sediment in rivers is transported as either bedload (the coarser fragments which move close to the bed) or suspended load (finer fragments carried in the water). There is also a component carried as dissolved material.

For each grain size there is a specific velocity at which the grains start to move, called *entrainment velocity*. However the grains will continue to be transported even if the velocity falls below the entrainment velocity due to the reduced (or removed) friction between the grains and the river bed. Eventually the velocity will fall low enough for the grains to be deposited. This is shown by the Hjulström curve.

A river is continually picking up and dropping solid particles of rock and soil from its bed throughout its length. Where the river flow is fast, more particles are picked up than dropped. Where the river flow is slow, more particles are dropped than picked up. Areas where more particles are dropped are called alluvial or flood plains, and the dropped particles are called alluvium.

Even small streams make alluvial deposits, but it is in the flood plains and deltas of large rivers that large, geologically-significant alluvial deposits are found.

The amount of matter carried by a large river is enormous. The names of many rivers derive from the color that the transported matter gives the water. For example, the Huang He in China is literally translated "Yellow River", and the Mississippi River in the United States is also called "the Big Muddy". It has been estimated that the Mississippi River annually carries 406 million tons of sediment to the sea, the Yellow River 796 million tons, and the Po River in Italy 67 million tons.

12

Applied Geomorphology in Coastal-zone Management

COASTAL ZONE MANAGEMENT

Coastal zone management involves managing coastal areas to balance environmental, economic, human health, and human activities.

At the core of the Coastal Zone Management Act are two programs: the National Coastal Zone Management Program and the National Estuarine Research Reserve System (NERRS). The NERRS is a network of 29 protected areas, including the San Francisco Bay National Estuarine Research Reserve, established for long-term research, education and coastal stewardship.

The concept of coastal zone management is a relatively new one, emerging less than four decades ago from the need to tackle an array of interconnected problems associated with population growth and development along our nation's coasts.

The Coastal Zone Management Act (CZMA) was passed in 1972 and provided a formal structure to address the challenges of continued growth in coastal areas. Administered by NOAA, the CZMA recognizes that ensuring access to clean water and healthy ecosystems that support a vibrant coastal economy requires effectively integrating science, technology, and public policy. The goals of the CZMA are to "preserve, protect, develop, enhance, and restore where possible, the coastal resources." One program under the CZMA, the National Coastal Zone Management Program, encourages coastal states and territories to work in partnership with the federal government to design and enforce local programs consistent with the CZMA and accompanying regulations. Today, 34 of the 35 eligible coastal and Great Lakes states and territories have entered into the voluntary partnership.

As a result of the Coastal Zone Management Act and the success of its programs, coastal communities are equipped to better address continued economic development of the coastal zone while accounting for natural resource management. This will ensure the health and stability of the coast, both environmentally and economically, into the long-term future.

Coastal Zone Management- Purpose, Objective and Challenges

The Coastal zones are defined by the extent of territorial waters up to the high water mark. They are long, narrow features of mainland, islands and seas, generally forming the outer boundary of the coastal domain. Find out the concept, purpose, objectives and challenges of Coastal Zone Management for general awareness.

The Coastal zones are defined by the extent of territorial waters up to the high water mark. They are long, narrow features of mainland, islands and seas, generally forming the outer boundary of the coastal domain. The Coastal Zone Management (CZM) is a process of governance that consists of the legal and institutional framework necessary to ensure that development and management plans for coastal zones are integrated with environmental and social goals, and are developed with the participation of those affected.

Purpose of Coastal Zone Management

The goals of the Coastal Management (CZM) are to "preserve, protect, develop, enhance, and restore where possible, the coastal resources." The purposes of Coastal Zone Management are given below:

1. To maximize the benefits provided by the coastal zone
2. To minimize conflicts and harmful effects of activities upon each other, resources and the environment
3. To promote linkages between sectoral activities
4. To guide coastal area development in an ecologically sustainable fashion

Objective of Coastal Zone Management

The prime objective of CZM is to create balance between development needs and protection of natural resources which means if coastal ecosystems are manage through the guiding principles of sustainability then livelihoods of millions will be protected and their survival guaranteed.

"To preserve and protect the coastal zone from activities that may cause degradation of quality or loss of coastal land"

Challenges of Coastal Zone Management

The Coastal Zone Management is an interdisciplinary and inter-sectoral approach to problem definition and solutions in the coastal zone. Hence, faces numerous challenges which are given below:

1. Failure to appreciate the interconnections within coastal systems
2. Inadequate legislation and lack of enforcement
3. Limited understanding and experience in ICZM
4. Limited understanding of coastal and marine processes
5. Lack of trained personnel, relevant technologies and equipment

The Coastal Zone Management (CZM) is based on two sets of principle- firstly agreed international norms, which were set out in the Rio Declaration on Environment and Development; and secondly based on bio-physical nature of the coastal zone.

Coastal Zone Management Act (CZMA)

It was passed in 1972 and provided a formal structure to address the challenges of continued growth in coastal areas. It is administered by National Oceanic and Atmospheric Administration (NOAA). This act recognizes that ensuring access to clean water and healthy ecosystems that support a vibrant coastal economy requires effectively integrating science, technology, and public policy. The goals of the CZMA are to "preserve, protect, develop, enhance, and restore where possible, the coastal resources."

The specific character of coastal zones

A well-informed science-based coastal zone management strategy embedded in an adequate social, institutional and legal framework, can prevent many future coastal problems. This is now usually called ICZM, Integrated Coastal Zone Management. As far as the technical aspects are concerned, experienced coastal authorities are capable to overview most of the coastal engineering issues associated with the future developments of the coastal zone. However, ICZM requires a broader view of coastal issues. ICZM is a governance process for the coastal zone, which differs from usual territorial governance processes due to its specific characteristics:

1. The coastal zone has no fixed administrative boundary; it is defined by the environmental (physical, ecological) interaction processes between the land environment and the marine environment that evolve over time.
2. The coastal zone usually has no single entrusted government; coastal zone governance is an interplay of several local, regional and national institutions with different mandates and responsibilities.
3. The coastal zone environment (the physical and ecological state) is highly dynamic due to the interaction processes between the land environment and the marine environment.

4. The coastal zone offers important ecosystem services with a much wider geographical significance.
5. Settlements in low-lying coastal zones are very vulnerable to extreme climatic events and to the impact climate change.

Because of these particular characteristics many studies and experiments have been carried out for defining a coherent ICZM governance process for coastal zones. Studies and experiments for developing and implementing ICZM are listed in the category Integrated Coastal Zone Management.

Why it is difficult to put ICZM into practice

The problems with which coastal zones are confronted often have an insidious character, such as ongoing urban or touristic development, saline intrusion, decline of biodiversity or climate change and sea level rise/land subsidence. When problems are perceived as urgent the situation is often already irreversibly deteriorated. Who feels responsible for a well-balanced future-proof development of the coastal zone? At national level, coastal management responsibilities are divided over ministries serving different (and often competing) societal interests, such as public safety, infrastructure, nature, water, physical planning, housing, fisheries, agriculture, tourism, industry and cultural heritage. Local governments are often the only bodies responsible for weighing and integrating different interests. But the means to do this are limited, because often the situation is that:

- local authorities can be overruled by sectoral authorities on a higher governance level (region, state);
- local authorities have little or no staff with profound knowledge of the complex interactions that take place in the coastal zone;
- local authorities have limited financial resources for monitoring and assessment of the state of the coast and for restoration measures;
- local authorities have limited manpower and willingness to enforce regulations;

- The local interests that local authorities represent are often short-term interests.

Such obstacles to putting ICZM into practice can be overcome, at least in part, by delegating more powers to local authorities and by promoting public participation and involvement of civil society organizations (NGOs) in coastal policy processes. This may require a transformation of the existing governance culture. However, it is essential that ICZM is firmly anchored at a high political level and that institutions with widely accepted legitimacy play a leading role in the implementation.

ICZM implementation

The natural and social characteristics of different parts of the coastal zone can be highly diverse. Coastal zone policy can therefore be determined only partly at the national level. The primary focus at the national level is to establish a legal, institutional and administrative framework for integrated coastal zone management. Of crucial importance is the institutional embedding of the ICZM process. The institutional framework must provide the mandate and resources for the local implementation of ICZM. Implementation of ICZM requires that sufficient powers be delegated to local authorities. This can be a problem in countries with a strongly centralized governance culture.

The coastal zone is constantly evolving through natural and socio-economic processes. ICZM should therefore not consist of a one-time static plan or as a series of ad hoc actions, but must be shaped as a continuous process that goes through an established policy cycle according to the schedule:

Plan development => *Implementation* => *Monitoring* => *Evaluation* => *Plan revision* => *Implementation* => *Monitoring* => *Evaluation, etc.*

This cycle is implemented at both local and national level:

- At local level, detailed concrete plans are developed and carried out (after endorsement at the national level) in consultation with all local and national stakeholders;
- At national level, national objectives and targets are defined,

local plans are integrated in a national strategy and mandate and resources are allocated for implementation.

The cycle period should be adjusted to the rate at which developments take place in the coastal zone. A cycle of one year may be too short; a cycle of 5 years or 10 years can be more appropriate. The cycle period at the local level may be shorter than at the national level.

Monitoring and evaluation are essential process steps to determine which progress has been realized in the implementation of the ICZM plan. This is only possible if measurable indicators and quantitative targets have been defined for this purpose. Defining indicators and targets is a major component of the ICZM plan process. Various examples of ICZM indicators have been described in the literature. ICZM indicators proposed by Marti et al. (2007) are shown for illustration in Table 2. Successful implementation of ICZM is highly dependent on the definition and monitoring of adequate indicators and targets.

Institutional arrangements

In small countries with a homogeneous coastal zone (natural and socio-economical), local and national level can coincide. This will not be the case for large countries with very diverse coastal zones. Both at local and national level, sufficient expertise must be available to implement the ICZM policy cycle. At the national level, a coordinating ministry must be designated that steers the policy cycle at the national level. This coordinating ministry should have internally sufficient general ICZM expertise to coordinate coast-related policies of other ministries and be capable to mobilize expertise of public and private organizations on specific topics. At the local level, a department of local government is responsible for implementation. The staff of this local department should have expertise in the fields of planning, communication, organizing public participation, administrative and technical implementation aspects and cooperation with private parties.

Complex infrastructural works and works that transcend the local scale do not fit within the aforementioned scheme. The steering of design and implementation in these cases lies with institutions at national level that are mandated for this and possess the required expertise.

The issue of climate change, which has far-reaching consequences for low-lying coastal areas, in many cases exceeds the local scale in terms of extent and complexity. A national adaptation strategy will be leading here for the development of adaptation measures on a local scale.

Goals and Objectives of the ICZM Programme

The overall goal of the ICZM-type programme is to ensure optimum sustainable use of coastal natural resources, perpetual maintenance of high levels of biodiversity, and real conservation of critical habitats. Tangible objectives of ICZM include, for example, supporting fisheries, protecting the community from storm ravages, attracting tourists, promoting public health, maintaining yields from mangrove forests, preserving coral reefs. All these require coordinated community action for their accomplishment, a need that ICZM fulfils.

A major objective of ICZM is to coordinate the initiatives of the various coastal economic sectors (e.g., shipping, agriculture, fisheries) toward long-term optimal socio- economic outcomes, including resolution of conflicts between sectors and arranging beneficial trade-offs. Integrated in this way, a multi-sector approach could jointly guide the activities of the key economic sectors under an effective coastal planning and management system.

As an example of the above, both coastal tourism and fisheries depend to a large extent on a high level of environmental quality, including coastal water quality. Both sectors may be impacted by "spillover" effects such as pollution, loss of wildlife habitat and aesthetic degradation from uncontrolled oil and gas development. Therefore, these different sectors should be working through ICZM to extract safeguards from the oil industry.

In another example, fisheries may require port services similar to those which tourism depends on as well as an infrastructure system that supplies water, sanitation, transportation, and telecommunications. Therefore, plans for both tourism and fisheries should be integrated with those for transportation and public works through ICZM.

Resources management and environmental conservation, which provide the motivation for ICZM planning and management programmes, are not incompatible with economic growth. In fact, enhanced long-term economic development may be the overall driving force of ICZM. Its advocates must ensure that ICZM is not perceived to have only negative impacts on jobs, revenue, or foreign exchange. If this were the case, the programme would not be likely to survive even the initial planning stages.

Scope of the ICZM Programme

ICZM refers to a special type of governmental programme established for the purpose of conserving coastal resources or environments (Sorensen and McCreary, 1990) through control of development. Use of the term implies that the governmental unit administering the programme has distinguished a special coastal zone as a geographic area combining both ocean and terrestrial domains.

A coastal programme includes several types of resources and environments as well as the interests of the various economic sectors that are stakeholders in coastal resources. This latter aspect - the involvement of all parties of interest - is what distinguishes "Integrated" Coastal Zone Management". The sectors may include shipping, housing, manufacturing, fisheries, agriculture, electric power, transport, nature conservation, or tourism.

To accomplish its purposes, ICZM requires several national actions, including the following:

1. A policy commitment to support coastal resources management and environmental conservation
2. Achieving an understanding on resources and environmental objectives among the various coastal stakeholders
3. Establishment of a governmental office for coordination of coastal affairs

4. Initiation of a system for review of development projets, including environmental assessment
5. Accumulation of technical information
6. Design and development of effective planning and management programmes

The degree to which the above are accomplished governs both the level of ICZM programme integration (collaboration among numerous government and private economic sectors) and comprehensiveness (the scope of development control and resource management objectives) that can be achieved. High levels of both can be achieved in a full-scale ICZM programme, the fundamentals of which are addressed later in this report.

Where the above tableau appears too complex or daunting, the ICZM programme can be simplified. A simple ICZM programme needs to include only the following components:

1. Arrangements: Policy, goals, legal authorization and an enforcement mechanism
2. Coordination: A coordination office
3. Review: A project review/permit mechanism and environmental impact assessment procedure

To be successful, ICZM has to be a distinct process focusing on distinct issues. Its goals must be clear and unambiguous.

Extent of Jurisdiction

The ICZM programme is expected to focus most sharply on management of the physical development process using planning procedures and government regulations. Resource management will usually have been mandated to other agencies even before the ICZM programme is implemented. Because these mandates are not easily changed, management of fisheries and mangrove forest harvests will usually remain with existing resource agencies. Parks and reserves also may continue to be managed by the existing jurisdiction. Also pollution control may remain with the national agency responsible for environmental protection. ICZM should be seen as a multi-sectoral process created to improve development planning and resource conservation though integration and cooperation. It should not be seen as a substitute for uni-sectoral programmes such as tourism or maritime administration, nor as a substitute for coastal forestry or agriculture programmes.

Regarding fisheries, while ICZM programmes can help protect fish habitat, nursery areas, and water quality in fishery areas, they do not usually address fisheries operations such as controls on harvesting (e.g., quotas, sizes, closed seasons, gear type, etc.). Because these control functions typically involve only the fishery sector, they can be handled separately. As mentioned, ICZM usually addresses multi-sector concerns. However, the ICZM format can assist fisheries harvest management and such controls as closed areas and maintenance of traditional fishing practices. ICZM can support fisheries management particularly in "developing countries beset with overfishing problems and multiple resource use conflicts" (Chua Thia-Eng, pers. comm.).

Most of the management strategies discussed in this report require direct action of governments - regulatory actions or designation of coastal land or water areas as nature reserves or managed multiple-use areas.

The ICZM literature gives very little information about market-based incentives to the private sector such as:allocations, charging of user fees, tax incentives, and establishment of exclusive rights to resources (property rights). In the future, such approaches will probably be more prominent.

Because of the nature of the coastal zone and of enclosed seas resources and geographic circumstances, many of the main issues facing countries are of international scope. These problems may include sea transport, shared fisheries, cross-border pollution transport, multi-country endangered species.

Development projects are often financed by external sources. The major providers of investment capital to developing countries - World Bank, USAID, Inter-American Development Bank, Asia Development Bank, etc. - are slowly but surely taking the lead in a movement toward environmental responsibility.

Distinguishing ICZM

The natural resources of coastal areas are so different from their terrestrial counterparts as to require different and special forms of management. For example, coral reefs, beaches, coastal lagoons, submerged seagrass meadows, and intertidal mangrove forests have no counterparts in terrestrial resources. Threats to the productivity of these unique resource systems arise from development activities and their-side-effects, such as reef and beach mining, shoreline filling, marine construction, lagoon pollution, sedimentation, and other activities that are distinct from those on land. Therefore, a special and distinct management methodology is required, such as ICZM.

ICZM is interdisciplinary. It considers, coordinates, and integrates the interests of all appropriate economic sectors (Caddy, 1990). It is needed in order to cope with special conditions of coastal resource conservation and economic development. It is particularly useful in solving problems that exist between the various sectors; for example, in resolving conflicts among fisheries, tourism, oil and gas development, and public works where these sectors are all attempting to use the coastal zone simultaneously (Sorensen and McCreary, 1990).

The above needs are accomplished by managing the coastal area as a unit and by integrating the management process with all appropriate economic sectors (shipping, mining, housing, transportation, etc.). An integrated coastal zone management programme must address coastal waters and shorelands together in a single, unified programme. The programme must recognize coastal shorelands, lowlands, intertidal areas, lagoons, and open waters as a single interacting and indivisible resource unit that lies between the hinterlands and the open sea and whose future must be planned and managed as a unified whole.

This becomes a complicated undertaking because so many interests have a stake in how coastal resources are allocated. One potential simplification is zoning of the coastal area for advance geographic designation of permitted uses.

Development planners, particularly, must recognize that modification of the hinterlands has a high potential for adverse effects on resources of coastal zones, including enclosed seas (Clark, in press). Examples of major impacts are: land clearing and grading, agricultural pollution (runoff); siltation from eroded uplands; dams across major rivers; land filling to provide real estate for industry, housing, recreation, airports, and farmland; harbour improvement; mining/quarrying; and excessive cutting of mangroves for fuel.

Another distinguishing facet of ICZM is that it combines development management with resources management. General environmental conservation is a fundamental component. It includes some very important, but not so tangible objectives, such as nature protection and biodiversity conservation. The extent to which such environmental and bio-diversity conservation objectives should be included in any ICZM programme depends on the particular country's willingness to extend to nature the priorities usually given to economic yields, products, employment, income, and consumption. Recently, the nature motivation has been enhanced by the recent global awakening to the need for ecological conservation.

Boundaries of the Coastal Zone

The coastal area is the interface between the land and the sea and may extend inland and seaward to a variable extent, depending upon the objectives and needs of the particular programme. By virtually any set of criteria, the coastal zone is a linear band of land and water that straddles the coast - a "corridor" in planning parlance - which has a one-dimensional aspect. The second dimension (width from onshore to offshore) tends to be overshadowed by the linearity; thus people talk about being at the coast or on the coast, but but never in the coast.

The boundaries of the coastal zone depend on political, administrative, legal, ecological and pragmatic considerations because there is a broad array of possible coastal issues and because the zone can be affected by remote activities. A narrow coastal zone could be appropriate if its purpose were to manage only the shoreline and intertidal areas. If watershed issues are of concern, then an inland extension is necessary. Likewise, if the issues extend far seaward, then the Exclusive Economic Zone (EEZ) might be included. In the case of small island countries, the whole country might be defined as coastal zone.

From a strategic and political perspective, success in creating a national ICZM- type programme may come more easily if the zone for management is narrow than if it is wide. A mandate to manage a narrow strip of transitional area, comprised of mostly tidally influenced habitats, may win approval more easily than a broader zone including deeper parts of the sea and higher and drier parts of the shorelands. The zone should include, at least, the coastal lowlands, the intertidal areas, salt marshes, wetlands and beaches, and offshore features such as coral reefs and island habitats.

It is often the coastal common property resources of the “wetside” that are emphasized in ICZM rather than the private property of the “dryside”. These resources can be adequately managed by a programme that gives priority to the edge zone.

Sustainability

In most of the parts of the world, renewable coastal resources tend to be economically limited (Snedaker and Getter, 1985). Over time, the demand for a given resource will commonly exceed the supply (unless there is a corresponding increase in price), be it arable land, fresh water, wood or fish. Suistainable use management helps to ensure that renewable resources remain available to future generations in abundance (and at the lowest price).

Suistainable exploitation implies the wise use and careful management (conservation) of individual species and communities, together with the habitats and ecosystems on which they depend, so that their current or potential usefulness to people is not impaired.

Resources should be maintained so that the ability of a resource to renew itself is never jeopardized. Such management maintains biological potential and enhances long-term economic potential of renewable natural marine resources. The criterion for sustainable use is that the resource not be harvested, extracted or utilized in excess of the amount which can be regenerated. In essence, the resource is seen as a capital investment with an annual yield; it is therefore the yield that is utilized and not the capital investment which is the resource base. By sustaining the resource base, annual yields are assured in perpetuity (Snedaker and Getter, 1985).

Sustainability is the alternative to resource depletion caused by excessive exploitation for short-term profit. An economic corollary is that development projects should be designed to sustain themselves without continued monetary subsidy (from governments, international donors, etc)

For development and management purposes, the coastal zone is an economic entity, like farmlands, production forests, rangelands or cities. Coastal renewable resources should be managed to produce benefits on a long-term, sustainable, basis. Some fisheries are managed on the principle of "optimum sustained yield" (OSY) but other resources often are not. However, mangrove silviculture in Malaysia is an example of a planned sustainable yield programme.

Exclusive use of a particular coastal resource unit for a single economic purpose is discouraged by ICZM in favour of a balance of multiple uses whereby economic and social benefits are maximized - conservation and development then become compatible goals (if "conservation" is used in its classic sense).

The word "conservation" (in its English language version) came into use early in this century. It was used extensively in the USA by the administration of President Theodore Roosevelt (1901–1908) to mean, simply, "wise use" of natural resources or one might say, "the control of exploitation to ensure resource sustainability". In this classic sense "conservation" and "sustainable use" are nearly analogous.

Unfortunately, in a broader vernacular use, conservation is used interchangeably with such words as "protection" and "preservation" ; and " conservationist" is used interchangeably with "environmentalist" or "ecologist". Consequently, using conservation interchangeably with sustainable use could cause confusion. Therefore, (1) the term "resource management" is used as the agent for achieving sustainable use of resources and (2) "conservation" conjunctively as in environmental conservation (broad) or bio-diversity conservation (narrow).

Creating Protected Areas

Coastal and marine nature reserves should form a complementary and integral part of ICZM programmes. Such reserves are a necessary component of coastal resource stewardship. While the primary ICZM programme utilizes the regulatory powers of government to achieve resource conservation, by controlling private development activities and resolving potentially divisive conflicts among competing users of the coastal zone, this accomplishes only a part of the job.

Additional and more focused natural area conservation (the custodial approach) can be accomplished by exercising the proprietary rights (rights by virtue of ownership) of government through declaration of protected areas - resource reserves, natural areas, and/or national parks.

Protected areas complement and make possible other objectives of ICZM by conserving nursery areas for fisheries production, enhancing tourism revenues and recreational benefits, preserving wilderness values, promoting baseline scientific and management studies, and so forth.

At the same time, protected areas gain from the ICZM programme important protection from external impacts. Thus, ICZM and protected areas programmes can be mutually beneficial. Together they ensure the maintenance of a healthy resource base upon which to build sustainable development.

It is also possible for the resource users and private property owners to assist with coastal resources conservation through voluntary actions of their own. ICZM programmes can be organized to encourage private protection of natural areas through provision of incentives.

STATUS OF THE COASTAL ZONE MANAGEMENT

The coastal areas comprise 20% of the earth's surface and yet contain 50% of the entire human population. By the year 2025, the coastal population is expected to account for about 75% of the total population of the world. More than 70% of the world's megacities are located in coastal areas. Coastal ecosystems yield 90% of the global fisheries. Hence, it becomes truly important to conserve the coastal zones. The National Coastal Zone Management Program established under the Coastal Zone Management Act, 1972, takes into account the benefits of marine habitats to the Indian economy and encourages coastal states to collaborate with the federal government under developing programs to protect and improve the areas. With the aid of the coastal zone management act, coastal regions are better prepared to deal with problems that occur constantly. The Pacific and Atlantic coastlines are often threatened by external causes, according to the National Ocean and Atmospheric Administration. Drilling industry faces a high risk of oil spill into the oceans – the damage caused by spills is massive and very expensive. Yet all the coastal risks are not the product of human activity. Natural hazards like erosion, harmful algal blooms, big storms, flooding, tsunamis, and sea-level rise cause even greater harm, and half the Atlantic coast-living nation is well aware of that fact.

Coastal zone management programs pay a great deal of attention to forecasting hurricanes and tsunamis and mitigating disruption to coastal areas. Constantly rising sea level is certainly not good news for the coastal community, meaning the National Oceanic and Atmospheric Administration (NOAA) still has several initiatives to introduce. Natural resource security is particularly essential for the coastal areas. The leisure potential and the coastal habitats have been impaired to some degree because of intensive industrial and human activities.

The coastal zone is a complex and frequently volatile environment that is affected by the inland ecosystems (e.g. land use, drainage and groundwater outflow), the atmosphere and the ocean. The systematic integration of several fields of science is therefore important in order to understand these various factors and to establish a comprehensive strategy for coping with natural and anthropogenic factors on the coastal zone.

Need for coastal zone management

The coastal zone constitutes a region of transition between maritime and territorial regions. It is estimated that approximately 4800 billion tons of household waste and 65 million tons of solid waste is dumped annually in the sea. Due to these ongoing attacks on the mangroves, and coral reefs, fish breeding is constantly decreasing and affecting the livelihood of 200 million people living along our country's 7517 kilometer long coastline.

In February 1991, the Union government released a notification requiring all coastal states and union territories to prepare their coastal zone management plans within one year from the date of notification. The plans were not ready in time so in April 1996, the Supreme Court ordered the states in question to prepare the Coastal zone management plans by the end of September 1996.

In compliance with the direction of the court, all the states and territories of the Union concerned submitted their proposals in due time, and they were approved by the Central Government with certain conditions and amendments. Coastal states and union territories are now expected to control construction activities in compliance with the terms of the Coastal zone management plan approved notification. It is high time that this notification must be withdrawn in order to protect vulnerable coastal habitats because it not only is harming the creatures at the sea coast but is also affecting the lifestyles of the individuals living by the seacoast, as mentioned in the case of *The Kerala State State Coastal… vs The State Of Kerala Maradu*.

Propping up of a full city with a multi-million population, although temporary, needs to have a proper environmental impact assessment and a management plan in place, particularly if it is being constructed on highly ecologically sensitive land, including a river bed and its immediate floodplains. Due to all the reasons mentioned above, the need of the hour is the promulgation and enactment of a new Coastal Zone Protection Act with the specific definition of different areas, after due consultation with the fishing communities, stakeholders, scientists and the appropriate government. Improving coastal resource efficiency and utilization is critical.

The Indian Ocean includes large amounts of minerals including products including cobalt, copper, manganese and rare earth. These minerals are important for the manufacture of smartphones, laptops and components for the electronic industry, etc. Seawater also contains commercially useful salts including gypsum and common salt. Gypsum is useful in various industries. This would help Make in India initiative.

Coastal management includes habitat conservation initiatives such as afforestation/shelter beds for mangroves, aquatic plant regeneration, sacred groves eco-restoration, etc. This is vital for the protection of coastal communities and helps in flood prevention. Sustainable development: Indian Ocean marine resources can act as a backbone of economic growth for India and can support India to become a $5 trillion economy by 2022. Blue economy, through sustainable use of oceans, has great potential for boosting economic growth.

Coastal management will contribute to community building and inclusive governance for the adoption and application of comprehensive approaches to coastal management. This would assist in local people's involvement leading to good governance which is critical for sustainable and inclusive development. Coastal management includes the creation of infrastructure for tourism, restoration and recharge of water bodies, beach cleaning and development, and other small infrastructure facilities. This is important to promote tourism in an environment-friendly way. It will provide many with jobs and better livelihoods. Inclusive development can improve. For example, improving the utilization of fishery resources will provide a livelihood for many.

Initiatives to improve living conditions, such as climate tolerant or salinity resistant agriculture, water harvesting and recharging/

storage, infrastructure and facilities to promote eco-tourism, community-based small-scale mariculture, seaweed cultivation, aquaponics etc. would do value addition to other livelihood activities. It is projected that marine fisheries resources along the Indian coastline have an annual harvestable capacity of 4.4 million metric tonnes.

Thus coastal management will give the fisheries sector a boost through infrastructure growth. Petroleum and gas hydrates are the principal energy resources available in the Indian Ocean. The oil extracted from offshore regions comprises mainly petroleum products.

Gas hydrates are unusually compact chemical structures made of water and natural gas. Along with this tidal energy is also important that would provide electricity to locals. This will result in food security through fisheries and other marine food resources. This will also help to popularize the problem of malnutrition in India because fish are a good source of nutrition. Coastal management includes infrastructure growth along the coast. Coasts are a major gateway to trade. Better connectivity in the area will reduce transport costs dramatically and reduce logistic inefficiencies.

Coastal zone regulation notification

Coastal zone was always on the forefront of civilisations and by far has been the most exploited geomorphic unit of earth. As we have already seen that, the coastal world is under increasing pressure, there is an urgent need to implement and conserve the zone. Coastal Areas of creeks, bays, oceans, rivers, and backwaters that are influenced by tidal activity of up to 500 meters from the High Tide Line and the land between the Low Tide Line and the High Tide Line are Coastal Control Zones (CRZ). The Government of India notified Coastal Regulation Zones (CRZ) for the first time.

Coastal zones mentioned under the notification:

As per the notification, coastal areas have been classified into four categories as CRZ-1, CRZ-2, CRZ-3 and CRZ-4.

CRZ-1

These are environmentally sensitive areas which are important for the protection of the coastal ecosystem. These include national parks/marine parks, sanctuaries, forest reserves, areas for birds, mangroves, and coral/coral reefs. Such areas are situated between the high and low tide lines.

CRZ-2

This includes areas which have already grown before the coastline. Construction of unauthorised structures is prohibited in this zone.

CRZ-3

Rural and urban settlements are included under CRZ-3 which are largely undisturbed and do not belong to the first two groups. Only limited agricultural or some public facilities-related activities are permitted in this region. This covers non-substantially built-up areas within city boundaries or in legally designated urban areas.

CRZ-4

These areas include Lakshadweep's coastal coasts, the Andaman and Nicobar Islands and several other small islands, with the exception of CRZ-I, CRZ-II or CRZ-III. These areas exist up to the territorial limits within the aquatic zone. In this zone activities such as fishing and other related services are permitted. Releasing of solid waste on such land is forbidden. In addition to these zones, in order to conserve the resources by controlling their depletion, the government of India has also enacted the Environmental Protection Act, 1986 making it easier to conserve the water bodies. Areas requiring special protection

Management of the coastal zone is necessary in order to grow and use the coasts sustainably for sustainable growth. What's required is the use of technology, community engagement and elimination of bottlenecks such as lack of communication at different levels. Industries are drawn to the coastal areas as they:

1) Benefit from exposure to low-cost marine and inland transport systems;
2) Use seawater for refining or cooling purposes;
3) Deal with marine transport; and
4) Directly rely on the marine environment for raw materials. Coastal heavy industry waste waters have a significant detrimental effect on coastal habitats.

Such impacts vary from relatively minor disruptions to major disruptions. But the introduction of current, affordable, waste management technologies will remove much of the industrial emissions. Many estuaries, lagoons, and bays that are now polluted can also be restored successfully if the wastes entering them are adequately treated.

The economic importance of the fisheries is still overlooked by planners. With nearly 100 million tons of fish per annum, the sea produces more protein worldwide than combined beef and mutton. Fisheries also provide livelihoods for fishermen and their families and those in the fishing industry, including boat builders, trappers and net producers, packagers, distributors and retailers- all of which lead to social change, cultural, economic, and political stability in the coastal areas. A strong domestic fishery promotes self-sufficiency and reduces the outflow of foreign exchange. Also, profitable coastal fisheries reduce rural population migrations to already overcrowded cities.

For half a century, fisheries management has been a major concern of national and regional fisheries authorities, and the UN Convention on the Law of the Sea (UNCLOS) 1982, offers a basis for this. This concern is overshadowed by the problems resulting from the loss or destruction of biodiversity in the estuarine, riverine, and coastal seas. For the vast majority of species that spend their youngest periods in near coastal, estuarine-brackish, or fresh waters, the survival of marine fisheries is now endangered by coastal depletion.

In reality, sustaining the present production of seafood and further development, through improved fishing or aquaculture, would be difficult in the absence of coastal and enclosed environmental management.

Conclusion

Management of the coastal zone depends on the knowledge available on different aspects of coastal ecosystems, coastal processes, natural hazards and their impacts, quality of water and living resources. Successful management practices rely on the awareness and adequate response from the government agencies concerned. Loss of coastal marine areas is growing crucial environmental resources. Fish and shell populations could fall, and the shoreline could be destabilized.

In general, the relationship between wetland safety and fish stocks is linear. Wetlands, i.e. the wave attenuation service which protects coasts from storms and tsunamis, provide the ecological services.

The relationship between this service area and the wetlands is non-linear. These partnerships have management implications and should be taken into consideration.

We need to model the impacts of those processes and devise adaptation and mitigation strategies for the coastal zone's sustainable development. Geospatial information technology may make a significant contribution to the creation of these models.Integrated coastal zone management

Integrated coastal zone management means the integration of all aspects of the coastal zone; this includes environmentally, socially, culturally politically and economically to meet a sustainable balance all around. Sustainability is the goal to allow development yet protect the environment in which we develop. Coastal zones are fragile and do not do well with change so it is important to acquire sustainable development. The integration from all views will entitle a holistic view for the best implementation and management of that country, region and local scales.

The five types of integration include integration among sectors, integration between land and water elements of the coastal zone, integration amount levels of government, integration between nations and integration among disciplines are all essential to meet the needs for implementation. Management practices include

1. maintaining the functional integrity of the coastal resource systems,without disrupting the environment
2. reducing resource-use conflicts, by making sure resources are used adequately and sustainably,
3. maintaining the health of the environment, which means to protect the ecosystems and natural cycle,
4. facilitating the progress of multisectoral development, which means allowing developers to develop within standards.

These four management practices should be based on a bottom-up approach, meaning the approach starts from a local level which is more intimate to the specific environment of that area. After assessment from the local level, the state and federal input can be implemented. The bottom-up approach is key for protecting the local environments because there is a diversity of environments that have specific needs all over the world.

Adaptive management

Adaptive management is another practice of development adaptation with the environment. Resources are the major factor when managing adaptively to a certain environment to accommodate all the needs of development and ecosystems. Strategies used must be flexible by either passive or active adaptive management include these key features:

- AIterative decision-making (evaluating results and adjusting actions on the basis of what has been learned)
- Feedback between monitoring and decisions (learning process)
- Explicit characterization of system uncertainty through multi-model inference (experimentation)
- Embracing risk and uncertainty as a way of building understanding (trial and error)

To achieve adaptive management is testing the assumptions to achieve a desired outcome, such as trial and error, find the best known strategy then monitoring it to adapt to the environment, and learning the outcomes of success and failures of a project.

Mitigation

The purpose of mitigation is not only to minimize the loss of property damage, but minimize environmental damages due to development. To avoid impacts by not taking or limiting actions, to reduce or rectify impacts by rehabilitation or restoring the affected environments or instituting long-term maintenance operations and compensating for impacts by replacing or providing substitute environments for resources Structural mitigation is the current solution to eroding beaches and movement of sand is the use of engineered structures along the coast have been short lived and are only an illusion of safety to the public that result in long term damage of the coastline.

Structural management deals with the use of the following: groins which are man-made solution to longshore current movements up and down the coast.

The use of groins are efficient to some extent yet cause erosion and sand build up father down the beaches. Bulkheads are man-made structures that help protect the homes built along the coast and other bodies of water that actually induce erosion in the long run.

Jetties are structures built to protect sand movement into the inlets where boats for fishing and recreation move through. The use of nonstructural mitigation is the practice of using organic and soft structures for solutions to protect against coastal hazards. These include: artificial dunes, which are used to create dunes that have been either developed on or eroded. There needs to be at least two lines of dunes before any development can occur. Beach Nourishment is a major source of nonstructural mitigation to ensure that beaches are present for the communities and for the protection of the coastline. Vegetation is a key factor when protecting from erosion, specifically for to help stabilize dune erosion.

Coastal environments

There are many different types of environments along the coasts of the United States with very diverse features that affect, influence, and mold the near-shore processes that are involved. Understanding these ecosystems and environments can further advance the mitigating techniques and policy-making efforts against natural and man-made coastal hazards in these vulnerable areas.

The five most common types of coastal zones range from the northern ice-pushing, mountainous coastline of Alaska and Maine, the barrier island coasts facing the Atlantic, the steep, cliff-back headlands along the pacific coast, the marginal-sea type coastline of the Gulf region, and the coral reef coasts bordering Southern Florida and Hawaii.

Ice-pushing/mountainous coastline

These coastal regions along the northernmost part of the nation were affected predominantly by, along with the rest of the Pacific Coast, continuous tectonic activity, forming a very long, irregular, ridged, steep and mostly mountainous coastline. These environments are heavily occupied with permafrost and glaciers, which are the two major conditions affecting Alaska's Coastal Development.

Barrier island coastline

Barrier islands are a land form system that consists of fairly narrow strips of sand running parallel to the mainland and play a significant role in mitigating storm surges and oceans swells as natural storm events occur. The morphology of the various types and sizes of barrier islands depend on the wave energy, tidal range, basement controls, and sea level trends. The islands create multiple unique environments of wetland systems including marshes, estuaries, and lagoons.

Steep, cliff-backing abrasion coastline

The coastline along the western part of the nation consists of very steep, cliffed rock formations generally with vegetative slopes descending down and a fringing beach below. The various sedimentary, metamorphic, and volcanic rock formations assembled along a tectonically disturbed environment, all with altering resistances running perpendicular, cause the ridged, extensive stretch of uplifted cliffs that form the peninsulas, lagoons, and valleys.

Marginal-sea type coastline

The southern banks of the United States border the Gulf of Mexico, intersecting numerous rivers, forming many inlets bays, and lagoons along its coast, consisting of vast areas of marsh and wetlands. This region of landform is prone to natural disasters yet highly and continuously developed, with man-made structures attaining to water flow and control.

Coral reef coastline

Coral reefs are located off the shores of the southern Florida and Hawaii consisting of rough and complex natural structures along the bottom of the ocean floor with extremely diverse ecosystems, absorbing up to ninety percent of the energy dissipated from wind-generated waves.

This process is a significant buffer for the inner-lying coastlines, naturally protecting and minimizing the impact of storm surge and direct wave damage. Because of the highly diverse ecosystems, these coral reefs not only provide for the shoreline protection, but also deliver an abundant amount of services to fisheries and tourism, increasing its economic value.

Coastal Geography

The current trend does not mean that the Japanese automatically have a right to equidistance as the basis for these boundaries. Equidistance is certainly highly relevant and a basis for consideration, but, as the recent *Jan Mayen* and *St. Pierre and Miquelon* decisions illustrate, other considerations often result in delimitations different from equidistant lines. Thus, the proportionality analysis of the lengths of the relevant coastlines, the distance from the shoreline of the mainland and other considerations may cause the maritime boundary to be located closer to the Japanese islands than the equidistant line where these islands are small and far from the main Japanese islands, particularly where they face longer coastlines of other states.

The contemporary trend also means that claims based largely on natural prolongation, geology and geomorphology will not carry much weight. Equidistance, adjusted by coastline proportionality, and access to currently exploited resources will play important roles. Chinese claims based on the former considerations with regard to the South China Sea, the Yellow Sea and the East China Sea appear to have weakened as the law has evolved. Claims by Vietnam in the Gulf of Tonkin, and perhaps those of the Philippines and Vietnam in the South China Sea (especially regarding the Spratly Islands) may also have weakened.

In the Gulf of Tonkin, Vietnam seeks to limit the maritime boundary between it and China to a line allegedly established in 1887. This line runs due south from the ChineseVietnamese land boundary on the eastern side of the gulf, close to the large Chinese island of Hainan. It does not reflect the coastal front of Hainan and is located east of the equidistant line. The legal force of the 1887 line for maritime boundary delimitation purposes is highly questionable, but the matter must be resolved. If the 1887 line is not found to establish that boundary or historic rights, the geography of the coastline will become the significant factor under modern law. This argues for delimiting the boundary west of the 1887 line, reflecting the coastlines of the large Chinese island to the east and the Vietnamese mainland to the west. Such a boundary could approach equidistance.

In the Yellow Sea, maritime boundaries between the two Koreas and China have not been established. China has asserted as the basis for the maritime boundary in this area the limit of the silt deposited in the sea by the Yellow River. This silt line is located roughly two-thirds of the distance across the Yellow Sea towards the Korean Peninsula. China also claims fisheries zones in the Yellow Sea. While proof of historic exclusive fisheries may have some force in maritime boundary delimitation, the silt-line claim appears to be based on theories of natural prolongation and geomorphology. Such arguments have attracted little contemporary support, notwithstanding the language of the *North Sea Continental Shelf Cases*.

Moreover, delimitation based on coastal geography appears to be relatively simple, in the absence of complications arising from offshore islands distant from the mainland coasts or known significant geological or geomorphological features. The area may be characterised as essentially irregular upside-down Vwith equally long facing mainland coastlines. The maritime boundary is likely to approach equidistance if current practices are followed. The South Korean western islands (PaengnyOng-do and islands to the south) may create some complications in delimiting the maritime boundary between North and South Korea, a situation that would militate for some enclaving of the islands. A more complex picture is presented in the south, where offshore islands and disputed Chinese, Korean and Japanese claims come into play.

The serious territorial disputes over the islands in the South China Sea make it impossible to forecast the probable location of the maritime boundaries there. Certainly, sovereignty over the islands must be settled first. If one state prevails, the total effect of the islands may have a substantial impact on the overall location of the maritime boundaries. On the other hand, in the likely case that the islands are divided among several states, many complex minor maritime boundary lines will be required to divide the zones thus generated. Overall, however, the greater division of the South China Sea may be largely unaffected by the islands. The driving force will be the mainland coastlines.

The international maritime boundaries established by agreement and third-party dispute settlements demonstrate that small features distant from the mainland shore usually have a limited impact on the overall maritime delimitation within the 200-nautical-mile EEZ as a result of the application of various techniques for enclaving and discounting these features. Substantial islands far from shore may attract significant zones, but minor features may very well attract little area beyond the territorial sea. Many of the rocks, reefs, islets and islands in the areas discussed above are found in the latter category. Islands near the mainlands, however, can have augmenting effects. In general, the fundamental boundary delimitations for this area, if they are realised, will tend to enclave or ignore most of the islands and approach an equidistant line generated from the most substantial features. Many of the bordering states are opposite states, a circumstance that especially attracts resort to equidistance. Laterally related states will attract solutions that avoid cutoff effects (e.g., Brunei/ Malaysia).

Coastal geography appears to be the primary consideration in maritime delimitations. Connected to that consideration is a strong preference for delimiting maritime boundaries so that all the littoral states have access to the middle of the maritime areas in question. This approach was initially articulated by the first Judgement on maritime boundaries by the International Court of Justice in the *North Sea Continental Shelf Cases*. The Court found that a coastal state located within a concavity suffered a disadvantage from the convergence of equidistant lines that did not permit it to realise its equitable portion of the maritime zones involved. The result in that case was an agreement by the states on the North Sea to attribute areas to the state within the concavity that would give it access to the region towards the middle of the southern area of the North Sea. This regional view, based on equity, was even more explicitly endorsed by the tribunal in the *Guinea/Guinea-Bissau* arbitration, which took into account the entire regional configuration in order to produce an equitable result. The Court in the *Libya/Malta* case also considered the regional perspective in reaching its delimitation, as did the court of arbitration in the *Anglo-French* award. These cases establish that the regional perspective should be considered in maritime boundary delimitations.

This principle argues for providing each coastal state, where possible, with maritime boundaries that approach the high seas or the middle of the body of water in question. One of the interests to be protected is navigational and overflight rights, without the need to consider the interests or wishes of neighbouring states. In that regard, the North Sea solution gave Germany access towards the middle of the southern portion of the North Sea. In the *Anglo-French* case, this interest was considered to be generally relevant, but it was not a problem because of the favourable geographic circumstances.

In *St. Pierre and Miquelon*, the tribunal constructed a narrow 200-nautical-mile corridor that runs to the southernmost area within the tribunal's competence. This corridor can only be justified on the basis of providing access in the direction of the open sea. In the *Gulf of Fonseca* Judgement, the ICJ did not establish a corridor within the gulf, but, even more significantly, it confirmed Honduras's rights in the gulf including those of navigation and overflight.

In addition, it assured to Honduras the right to share in the area beyond the closing line of the gulf to the limits of national jurisdiction in the Pacific Ocean. A similar solution was employed by the ICJ in awarding the small island of Jan Mayen (Norway) a maritime boundary that approached the middle of the waters between it and Greenland (Denmark) despite Greenland's very long coastline in the area and Norway's access to the sea on the opposite side of Jan Mayen. These cases support the view that a state's access towards the open sea or the middle of the relevant body of water is a significant factor in delimiting maritime boundaries.

A review of the regional maps published in *International Maritime Boundaries* illustrates that in virtually all established maritime boundary delimitations (approximately 140), coastal states have not been cut off from access either to the open sea or to areas near the centre of the body of water on which their coastlines are located. This result is particularly apparent in the Caribbean Sea, the Persian Gulf, the Mediterranean Sea, the North Sea, the Baltic Sea, and the waters off Africa and South America. The African and South American situations are especially noteworthy because several of the coastal states agreed to maritime boundaries based on uniform cardinal directions and thus avoided cutoff effects on neighbouring states.

There are, however, a few exceptions to this practice. The maritime boundaries of Panama, albeit providing it with a fairly large maritime area, create a complete enclosure short of the middle of the bodies of water off both its coasts. In the Andaman Sea, the Indian islands of Andaman and Nicobar were given effect that substantially cut off similar access for Burma and Thailand. Nevertheless, these situations appear to be exceptions to the general practice in maritime boundary agreements.

Zones not Permitted by the Los Convention

Settlement of the maritime boundaries in the East Asian area may have to take up the unusual maritime jurisdiction claims of the littoral states. The LOS Convention and the related general international law permit coastal states to claim internal or archipelagic waters delimited by the normal baselines and closing lines, historic bay closing lines, systems of straight baselines and archipelagic baselines; 12-nautical-mile territorial seas; 12-nautical-mile contiguous zones beyond the territorial sea; 200-nautical-mile EEZs measured from the baseline for the territorial sea; continental shelves; and — under certain circumstances not present here — continental shelves beyond the EEZ. No other zones of jurisdiction are permitted. The Convention purports to be comprehensive and may be considered general international law in this regard.

In the Central East Asian area, other maritime jurisdiction claims appear to be common. Thus, North Korea claims a special 50-nautical-mile military zone in the Sea of Japan drawn seaward from an unprecedented baseline not permitted by any of the law of the sea conventions. China claims certain military exclusion zones and special fishery zones in the East China Sea. China and South Korea claim that prior notification is required for the transit of foreign warships in their territorial seas despite the absence of any language in the articles on innocent passage and transit passage of the LOS Convention supporting such claims. Similarly, Japan claims to exclude vessels carrying nuclear weapons from transiting its territorial sea, notwithstanding these provisions.

Such special claims and zones certainly arose out of the controversial relations among the neighbouring states in the region. In their bilateral relations, however, the coastal states may find that these unusual, if not unsanctioned, claims under the general international law of the sea do not present obstacles to maritime boundary delimitation. None of these claims violate jus cogens norms. Thus, these states may enter into international agreements that establish maritime boundaries accepting these unusual claims in their bilateral relations. For example, the USSR was apparently willing to acquiesce in the unprecedented claims of North Korea when it agreed to their maritime boundaries in the area.

The quid pro quo may have been North Korean acquiescence in the Soviet historic bay claim over Peter the Great Bay. Other unusual claims may pertain more directly to disputed boundaries, such as the Chinese special zones in the East China Sea, which appear to run into the eastern portions of the sea claimed by Japan and South Korea. The establishment of a definitive traditional maritime boundary would require the resolution of these Chinese claims, although joint arrangements or special rights might be included in the settlement. Of course, such settlements may not affect third-state navigation and overflight rights and the underlying disputes may therefore have to be resolved by other means.

Future Settlements

The entry into force of the Law of the Sea Convention may encourage the states in the region to settle their maritime boundaries. The Convention provides the requisite codified law and thus eliminates some grounds for dispute. It also clarifies some new law that bears on maritime boundary delimitation. Furthermore, it establishes a comprehensive system for resolving disputes involving the law of the sea.

This system permits states to invoke a variety of techniques, including negotiation, conciliation, arbitration and adjudication, to arrive at boundary settlements. The ready availability of conciliation procedures and the participation of specialists may encourage Central East Asian states to accept delimitation as well as other solutions. Since these states generally do not favour litigation, good offices and conciliation may play important roles in resolving.

The Convention does provide for compulsory binding settlement of disputes arising out of the application or interpretation of its provisions. The choice of procedure is left to the states concerned and, in the case of maritime boundary disputes, permits states to exclude certain compulsory and binding dispute settlement procedures. This is complex and somewhat ambiguous. It is clear, however, that unless a state submits a contrary declaration in writing at the time of signature, ratification or accession, maritime boundary disputes involving that state will be subject to the compulsory binding dispute settlement procedures established by the Convention. If such a declaration is deposited, it will not fully exclude disputes arising subsequent to entry into force of the Convention (November 16, 1994) from the dispute settlement system when negotiations have been exhausted. Those new disputes will remain subject to compulsory conciliation.

Furthermore, the Convention states that "such submission" excludes disputes "concerning sovereignty or other rights over continental or insular land territory." "Such submission" could be interpreted broadly to refer to all compulsory binding dispute settlement procedures regarding maritime boundaries. But its meaning is narrower. It appears in the context of a unilateral declaration excluding maritime boundary disputes from the normal dispute settlement provisions of the Convention (Part XV), and it follows directly upon language regarding submission to conciliation of newly arising disputes.

Thus, the text should be interpreted to mean that this special conciliation procedure may not be used for land territory issues. Absent a declaration, however, boundary disputes are properly subject to compulsory binding dispute settlement and the other dispute settlement procedures of the Convention. The relevant forum is not limited by this from addressing all issues necessary to determine the maritime boundary, including territorial sovereignty if it is germane. Furthermore, of course, states may specifically agree to submit a maritime boundary dispute to a different third-party forum for resolution, with or without qualifications. It remains an open question whether any of the disputes in the Central East Asian area will be resolved through binding third-party dispute settlement. Conciliation under the LOS Convention or outside the Convention may be the preferred route if negotiations fail. Conciliation may produce results that approximate a binding settlement since the process and consequent report may put the disputing states under strong mutual and international pressure to accept those results or otherwise settle the matter.

On the other hand, if a state elects to exclude maritime boundary disputes from compulsory binding settlement, compulsory conciliation applies only to newly arising maritime boundary disputes, which appears to rule out all of the disputes discussed in this chapter.

Since World War II, relations between certain of the littoral states in the region have been difficult. These difficulties have lessened the prospects for settlement of the maritime boundary disputes. Thus, North and South Korea have fought a war and enmity remains. China and Taiwan have entertained similarly difficult relations. Many nations in the region consider Japan to have been the aggressor during World War II and treat it warily. Their caution has increased as Japan has regained its economic and military strength. Vietnam was divided for a long time, and war consumed the energies of its two parts. Finally, some of these states feared that the Soviet Union harboured aggressive tendencies in the region.

This situation may now be changing as a result of the demise of the USSR and Russia's assumption of the Soviet territories in Central East Asia. In addition, progress is being made in bringing North Korea out of isolation, and a rapprochement between North and South Korea is a possibility. China and Taiwan have recently exchanged diplomatic visits and relations may move towards regularisation. Russia is not viewed as having maintained the Soviet aggressive tendencies and settlement of the dispute over the sovereignty of the Kuril Islands is under consideration. China has become increasingly integrated into the international community, but it has also begun to project strong regional aspirations. Many of these developments may bode well for discussions and, perhaps, ultimate resolution of the maritime boundary disputes. For several years Indonesia has led efforts to settle the disputes regarding the South China Sea without resolving the underlying territorial questions or establishing maritime boundaries. Non-boundary-based solutions may be the most promising in light of the difficult sovereignty and boundary issues involved. The early resolution of these disputes may succeed in preventing costly international crises. The South China Sea is a major shipping route, especially for supertankers moving from the Persian Gulf to Japan and Korea, container ships sailing between Hong Kong and Singapore, and military vessels. The extreme weather conditions, as well as the many shoals, pose a constant risk of major pollution incidents. Other environmental problems are also likely. Fishing by many nations in the area carries the risk of overexploitation and collapse of the fisheries.

Regional regimes to address these issues without prejudice to the underlying claims could be created. A variety of regional authorities could address these regional issues. They might vary with regard to the degree and nature of decision making, government control and enforcement roles. A possible first step might be to develop plans to respond effectively to potential environmental disasters. These arrangements might be addressed to common concerns and would be least likely to raise jurisdictional issues. They might even help to build trust among the states in the region. A second step might be to develop regional environmental standards implementing the pertinent provisions of the LOS Convention to assure that generally accepted international minimum standards are fully observed.

Higher environmental standards consistent with the Convention might be established within the littoral states' EEZs and for their flag vessels operating in the region, as well as with regard to the ocean dumping of materials shipped from their ports. Special standards might be adopted for particularly delicate areas of the EEZ, as permitted under the Convention. Once mutual cooperation on these matters is established, the region might move to a third step addressing fisheries management from a regional perspective by promoting a coordinated policy of conservation and effective enforcement procedures. The norms of the LOS Convention, the fact that these are sernienclosed seas and the impending United Nations Convention on Straddling Fish Stocks and Highly Migratory Fish Stocks should help to guide this effort.

Finally, cooperative arrangements for the development of seabed mineral resources could be addressed. Since all of these arrangements could be established without prejudice to the underlying territorial and maritime boundary issues, they might constitute optimal solutions that would defer the more politically charged issues to the indefinite future. Although the informal discussions promoted by Indonesia may be productive, recent reports are not encouraging. For a cooperative non-boundary solution to succeed, the participating states must have good relations that facilitate the mutual confidence necessary for effective working relationships. The current relations in the region do not bode well in this regard, but recent political developments may change this prognosis, and a phased process might help to eliminate the barriers.

The littoral states may have other incentives to reach settlements, especially before strong conflicts arise or crises occur. Recent maritime boundary practices and decisions tend to avoid upsetting long-standing resource development operations. Where many states are exploiting the living resources in an area, the solution may be to divide the resource among them. Time is on the side of the state pursuing development regardless of historic or traditional activities. If economic development of a marine resource takes place, it will be difficult to dislodge that state completely from the area by negotiation or third-party dispute settlement. Thus, the status quo may work against some of the states involved. They might wish to obtain boundary settlements before other states establish a new status quo of resource development. Recently, states in the region have unilaterally attempted to establish a new status quo that would be difficult to upset. Especially in the South China Sea, hydrocarbon leases have been issued, military garrisons have been built on the Spratlys, and military skirmishes have been fought.

Bibliography

Albritton, C. C. : *The Fabric of Geology: Reading*, Mass., Addison-Wesley Publishing Co., 1963.

Bhagwat, S.B. : *Foundation of Geology,* Global Vision Pub, Delhi, 2009.

Bhatta, B : *Remote Sensing and GIS,* Oxford University Press, Delhi, 2008.

Davies R. J., Cartwright J., Stewart S. A, Underhill, J. R., Lappin M.: *3D Seismic Technology: Application to the Exploration of Sedimentary Basins*, Geological Society, London, Memoir, 2003.

Elangovan K. : *GIS, Fundamentals, Applications and Implementations*, New India Publishing Agency, New Delhi 2006.

Gontz, A.M.: *Evolution of Seabed Pockmarks in Penobscot Bay, Maine*, University of Maine, Orono, 2011.

Hadfield, Peter : *Sixty Seconds that Will Change the World: The Coming Tokyo Earthquake*, Boston, C.E. Tutle, Co., 1992.

Jakucs, Laszlo: *Morphogenetics of Karst Regions: Variants of Karst Evolution,* Halstead Press, John Wiley and Sons, New York, 1977.

Jakucs, Laszlo: *Morphogenetics of Karst Regions: Variants of Karst Evolution,* Halstead Press, John Wiley and Sons, New York, 1977.

Kahn, Miriam B. : *Disaster Response and Planning for Libraries,* Chicago, American Library Association, 1998.

Kiefer, R. W. : *Remote Sensing and Image Interpretation*, New York, 1987.

Kramer, Steven L.: *Geotechnical Earthquake Engineering*, Prentice-Hall, 1996.

Kuiper, Gerard P. : *The Earth as a Planet*, Chicago, Univ. of Chicago Press, 1954.

Laskin, David : *Braving the Elements: The Stormy History of American Weather*, New York, Doubleday, 1996.

Lyell, C. : *Principles of Geology*, Boston, Little, Brown, 1853.

Mack, Pamela Etter : *Viewing the Earth: The Social Construction of the Landsat Satellite System*, Cambridge, MIT Press, 1900.

Mallik, T K : *Marine Geology : A Scenario Around Indian Coasts*, New Academic Pub, Delhi, 2008.

Mathur, S M : *Concise Glossary of Geology*, Scientific, Delhi, 2007.

Matt Duckham. : *GIS, a Computing Perspective*, Boca Raton, CRC Press, 2004.

Nicholls, A.: *Low Attainers and the Teaching of Geography*, Sheffield, Geographical Association, 1988.

Nierenberg, W. A.: *Encyclopedia of Earth System Science*, San Diego, California: Academic Press, 1992.

Okoshi, T. : *Three-dimensional Imaging Techniques*, Academic Press Inc., New York, 1976.

Park, RG: *Foundations of Structural Geology,* Routledge, 2004.

Ramasamy, S. : *Remote Sensing in Water Resources,* Rawat, 2005.

Rondinelli, D.A.: *Applied Methods of Regional Analysis: the Spatial Dimension of Development Policy*, Westview Press, Boulder and London, 1985.

Scheidegger, Adrian E.: *Morphotectonics*, Springer, Berlin, 2004.

Sebba, J.: *Geography for All*, London, David Fulton, 1996.

Sokhi, B.S. and S.M. Rashid : *Remote Sensing of Urban Environment,* Manak, Delhi, 1999.

Steven, A. : *Grand Canyon: Monument to Catastrophe,* Institute for Creation Research, London,1994.

Vergne, Jack: *Hard Rock Miner's Handbook*. Tempe/North Bay: McIntosh Engineering, 2003.

Viles, Heather: *Biogeomorphology*, Basil Blackwell, Oxford, 1988.

Wainwright, John; Mulligan, M.: *Environmental Modelling: Finding Simplicity in Complexity*, John Wiley and Sons Ltd, London, 2003.

Walter Leal: *Innovation, Education, and Communication for Sustainable Development*, New York, Peter Lang Publishing, 2006.

Wendy J. Gordon : *The Economics of Copyright, Developments in Research and Analysis* Edward Elgar, 2003.

Williams, Paul W.: *Catena Supplement 25,* Catena Verlag, Cremlingen-Destedt, Germany, 1993.

Yang, Xin-She: *Introductory Mathematics for Earth Scientists*, Dunedin Academic, Edinburgh, 2009.

Index

L

N

P

Q

R

S

T

V

❑❑❑